

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med* 2014;371:2488-98. DOI: [10.1056/NEJMoa1408617](https://doi.org/10.1056/NEJMoa1408617)

Supplementary Appendix

Table of Contents

Section	Page
Supplementary Methods	1
Table S1 Cohort descriptions and baseline characteristics	6
Table S2 List of hematopoietic genes and variants queried	9
Table S3 Called somatic variants in 160 hematopoietic genes	12
Table S4 List of non-hematopoietic genes and variants queried	21
Table S5 Called variants in non-hematopoietic genes	22
Table S6 Logistic regression for factors associated with mutations	23
Table S7 Logistic regression for factors associated with clonality by ancestry group	24
Table S8 Details on subjects that developed hematologic malignancies	25
Table S9 Logistic regression for factors associated with RDW \geq 14.5%	26
Table S10 Association of cytopenias with mutations	27
Table S11 Association of mutations with known versus unknown causes of anemia	28
Table S12 Risks associated with developing incident coronary heart disease and ischemic stroke using traditional risk factors and mutations	29
Figure S1 Risks associated with developing incident coronary heart disease using traditional risk factors as well as hsCRP, RDW, and mutations	30
Figure S2 Characteristics of DNMT3A variants	31
Figure S3 Co-mutation plot	32
Figure S4 Longitudinal assessment of mutations	33
Figure S5 Factors associated with mutations	34
Figure S6 Mutations by ethnic background	35
Figure S7 Blood counts for individuals with and without detectable mutations	36
Figure S8 Density plot for RDW	37
Figure S9 Kaplan-Meier Curves for overall survival by cohorts	38
Figure S10 Cause-specific mortality associated with mutations	39
Figure S11 Survival curves for individuals with clonal hematopoiesis	40
Figure S12 Association of somatic mutations with incident cardiovascular disease	41
GoT2D and T2D-GENES Consortia Information	42
SIGMA T2D Genetics Consortium Information	46

Supplementary Methods

Sample ascertainment and cohort descriptions

Subjects were ascertained as cases and controls for T2D from 22 cohorts in 3 consortia (Supplementary Table 1). Details on sample ascertainment for cases and controls from the most of the GoT2D (8 cohorts, 2,376 subjects) and SIGMA (4 cohorts, 3,435 subjects) consortia have been previously described^{1,2}. The other ascertained individuals were part of T2D-GENES (9,990 subjects), a consortium comprised of 10 population-based cohorts. Details on sample ascertainment can be found in Supplementary Table 1. The remaining 1,381 individuals were additional subjects in the Jackson Heart Study (JHS), a large population-based cohort of African-Americans in Jackson, Mississippi, who had given consent for genetic testing but were not in any previously sequenced cohorts³. Including the 1,027 subjects from JHS enrolled through T2D-GENES (513 with T2D), a total of 2,408 subjects from JHS were in this study. Since 3,400 subjects in JHS were consented for genetic studies, ~70% of consented subjects from JHS are represented in this study, with a modest overall enrichment for T2D.

Subjects for which age was not available (116 subjects) or with cell lines as the source DNA (492 subjects) were excluded, including all subjects from the Wellcome Trust Case Control Consortium (WTCCC).

Vital status for Finland-United States Investigation of NIDDM Genetics Study (FUSION), The Botnia Study (Botnia), Helsinki Siblings with Diabetes cohort (Helsinki_sib), and Scania Diabetes Register (Diabetes_reg) was ascertained from the Finnish or Swedish Hospital Death Registries. Subjects from Botnia, Helsinki_sib, and Diabetes_reg were pooled for survival analysis. Vital status for subjects from the Multiethnic Cohort (MEC) was ascertained from Center for Medicare Services (CMS) data. Vital status for subjects from the JHS was ascertained from vital records and annual follow-up interviews. For individuals lost to follow-up, if there was no death certificate, the individual was assumed to be alive. Vital status for non-diabetic Ashkenazis in the Longevity Genes Project (LGP) was ascertained from hospital death records and annual follow-up interviews.

Malignancy information for MEC was ascertained through linkage of the MEC with cancer registries of California and Hawaii. Malignancy information for JHS was ascertained from annual interview. Malignancy information for some subjects from FUSION and Botnia was available from the Finnish Hospital Discharge Register and Death Register, but not included because it was deemed to have significant ascertainment bias.

Standard blood cell indices (white blood cell count, hemoglobin, hematocrit, platelet count, and white blood cell differential) were available for most (but not all) subjects from JHS, LGP, Botnia, Malmo-sib, and Helsinki-sib. Information on red blood cell distribution width (RDW) was available on most subjects in JHS and LGP.

Data on cardiovascular outcomes for JHS was obtained from annual patient interview and adjudicated from hospital records. Data on cardiovascular outcomes for FUSION was obtained from Finnish Hospital Discharge and Death Registries. Coronary heart disease (CHD) included fatal and non-fatal myocardial infarctions as well as coronary revascularization procedures. For CHD analysis, those with prior CHD events were excluded. For ischemic stroke analysis, those with prior ischemic stroke were excluded. Lab data (blood pressure, body mass index, serum high density lipoprotein, serum total cholesterol, and high-sensitivity C-reactive protein) was obtained at the same time as blood collection for DNA.

Exome sequencing

DNA was obtained from individual cohorts and further processed at the Broad Institute of MIT and Harvard. DNA libraries were bar coded using the Illumina index read strategy, exon capture was performed using Agilent Sure-Select Human All Exon v2.0, and sequencing was performed by Illumina HiSeq2000. Sequence data were aligned by the Picard (<http://picard.sourceforge.net>) pipeline using reference genome hg19 with the BWA algorithm ⁴ and processed with the Genome Analysis Toolkit (GATK) to recalibrate base-quality scores and perform local realignment around known insertions and deletions (indels) ⁵. BAM files were then analyzed for single nucleotide variants using MuTect (<http://www.broadinstitute.org/cancer/cga/mutect>) with Oxo-G filtering (<http://www.broadinstitute.org/cancer/cga/dtoxog>) and for indels using Indelocator (<http://www.broadinstitute.org/cancer/cga/indelocator>), followed by annotation using Oncotator (<http://www.broadinstitute.org/cancer/cga/oncotator/>)⁶. All MuTect and Indelocator analyses were performed using the Firehose pipeline (<http://www.broadinstitute.org/cancer/cga/Firehose>) at the Broad Institute.

Variant calling

Cancer genome studies typically compare sequence from tumor and germline DNA, and define somatic mutations as the sequence variants present in tumor but not germline DNA. To circumvent the lack of matched tissue in this study, we defined a list of pathogenic variants reported in the literature and/or the Catalog of Somatic Mutations in Cancer (COSMIC, <http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/>) in human hematologic malignancies from 160 genes (see Supplementary Table S2). We specifically excluded genes known to be involved in hematologic malignancy that had a relatively high frequency of heterozygous loss-of-function germline mutations in the population (*NF1*, *SH2B3*, *BRCA1/2*). Identical frameshift variants that were seen 3 or more times from the same ancestry group were also excluded, unless such variants were previously reported as somatic. Frameshift and nonsense mutations were further excluded if they occurred in the first or last 10% of the gene open reading frame ⁷, unless mutations in those regions had been previously reported, (e.g. *DNMT3A* ⁸). We used minimum variant read counts of 3 for MuTect and 6 for

Indelocator. Python (<http://www.python.org>) scripts were used to parse mutation annotation format (MAF) files produced by MuTect and Indelocator for variants of interest.

To further confirm that we were detecting bona fide somatic mutations, we examined variants in 40 driver genes involved in non-hematologic malignancies⁹. We hypothesized that very few variants would be detected from these genes if our methodology had high specificity for real mutations. Using the same calling approach as with hematologic genes (Supplementary Table S4), we detected only 10 variants that were the same as those mutated in non-hematologic cancers, and most of these appeared to be rare germline polymorphisms as evidenced by allele fraction (Supplementary Table S5).

Targeted re-sequencing

Validation of variants discovered by whole exome sequencing was done with "Rapid Heme Panel" (RHP), a Laboratory Developed Test designed and validated at a CLIA-certified lab (Center for Advanced Molecular Diagnostics, Brigham and Women's Hospital). RHP uses TruSeq Custom Amplicon Kit (Illumina, Inc. San Diego, CA, USA) and contains 95 genes (50 for AML/MDS, 8 for MPN, 27 for ALL, and 10 others). For oncogenes, known mutation hotspots are targeted; and for tumor suppressor genes the entire coding sequence is analyzed. The average amplicon size is 250-bp and about 50% of the regions are covered on both strands. Library preparation was according to manufacturer's instruction and sequencing was 150 bp paired-end reads with MiSeq v2.2 chemistry. Raw data was analyzed with Illumina on-board Real-Time-Analysis (RTA v.2.4.60.8) software and MiSeq Reporter. The VCF files were filtered with a cutoff for any nucleotide position with 10 or more variant reads or with 5-9 variant reads (if allele frequency >33%) as well as a Q score greater than 30 and germline single nucleotide polymorphisms were removed by comparison to dbSNP database (NCBI Human Build 141). The filtered variant lists were manually reviewed and BAM file examined in Integrated Genome Viewer (IGV, Broad Institute).

For 13 subjects from JHS, DNA obtained from a peripheral blood sample collected 4 to 8 years after the original DNA was available for analysis. RHP was used as described above to assess VAF of the previously detected mutations at the second time point, and to assess for the acquisition of new mutations.

Genes: ABL1, ASXL1, ATM, BCL11B, BCOR, BCORL1, BRAF, BRCC3, CALR, CBL, CBLB, CD79B, CEBPA, CNOT3, CREBBP, CRLF2, CSF1R, CSF3R, CTCF, CTNNB1, CUX1, CXCR4, DNMT3A, DNMT3B, EED, EGFR, EP300, ETV6, EZH2, FANCL, FBXW7, FLT3, GATA1, GATA2, GATA3, GNAS, GNB1, IDH1, IDH2, IKZF1, IKZF2, IKZF3, IL7R, JAK1, JAK2, JAK3, KIT, KRAS, LUC7L2, MAP2K1, MEF2B, MPL, MYD88, NOTCH1, NOTCH2, NOTCH3, NPM1, NRAS, NT5C2, PAX5, PDGFRA, PDS5B, PHF6, PIGA, PIK3CA, PIM1, PRPF40B, PRPF8, PTEN, PTPN11, RAD21, RET, RIT1, RPL10, RUNX1, SETBP1, SETD2, SF1, SF3A1, SF3B1, SH2B3, SMC1A, SMC3, SRSF2, STAG2, STAT3, TET2, TLR2, TP53, U2AF1, U2AF2, WHSC1, WT1, XPO1, ZRSR2

Statistics and analysis plan

All statistical analyses were performed using R. Cox proportional hazards and Kaplan-Meier analysis was performed using the **survival** package (<http://cran.r-project.org/web/packages/survival/index.html>). Competing risks regression (CRR) was used to estimate hazard ratios for developing hematologic malignancy with death as the competing risk¹⁰. CRR and cumulative incidence analysis was performed using the **cmprsk** package (<http://cran.r-project.org/web/packages/cmprsk/index.html>). Fixed-effects meta-analysis using beta-coefficients for risk estimates from individual cohorts was used to provide summary hazard ratios across heterogeneous cohorts. Meta-analysis was performed using the **meta** package (<http://cran.r-project.org/web/packages/meta/index.html>).

Analyses for factors associated with clonal hematopoiesis were performed using logistic regression with the pre-determined variables age, sex, type 2 diabetes status, and ancestry.

Primary outcomes for clinical associations with clonal hematopoiesis were pre-determined to be all-cause mortality and hematologic malignancy, using Cox proportional hazards models and CRR, respectively. Association of mutations with blood counts was also a primary analysis. For hematologic malignancy outcomes, only events incident to the time of DNA collection were considered. Red cell distribution width was also included as a variable in the survival analysis because of its association with mutations and previous reports of its association with increased mortality. Other pre-determined covariates were age (categorical), sex, and T2D status.

An analysis of cause specific mortality revealed an increase in cardiovascular deaths. For this reason, we examined incident events of coronary heart disease and ischemic stroke using CRR and the pre-determined variables age, sex, type 2 diabetes status, body mass index, and systolic blood pressure. For some subjects, data was also available on smoking status, total cholesterol, and high-density lipoprotein. This was examined as a subgroup analysis in a separate regression (see Supplementary Table S12).

For some analyses, a cutoff of VAF at 0.10 was used. This value was chosen because it was close to the median VAF in the dataset and is roughly the lower limit of detection using Sanger DNA sequencing, and was thus used to designate large versus small clone size.

1. Flannick J, Beer NL, Bick AG, et al. Assessing the phenotypic effects in the general population of rare variants in genes for a dominant Mendelian form of diabetes. *Nature genetics* 2013;45:1380-5.

2. Consortium STD, Williams AL, Jacobs SB, et al. Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature* 2014;506:97-101.
3. Sempos CT, Bild DE, Manolio TA. Overview of the Jackson Heart Study: a study of cardiovascular diseases in African American men and women. *The American journal of the medical sciences* 1999;317:142-6.
4. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;25:1754-60.
5. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics* 2011;43:491-8.
6. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature biotechnology* 2013;31:213-9.
7. Ng PC, Levy S, Huang J, et al. Genetic variation in an individual human exome. *PLoS genetics* 2008;4:e1000160.
8. Walter MJ, Ding L, Shen D, et al. Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. *Leukemia* 2011;25:1153-8.
9. Lawrence MS, Stojanov P, Mermel CH, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 2014;505:495-501.
10. Fine JP and Gray, RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association* 1999;94:496-509.

Supplementary Table S1
Cohort descriptions and baseline characteristics

TOTAL	No T2D	T2D	T2D unknown	Clinical information available	Cohort descriptions and sample ascertainment	References
Overall (number with clone)	9303(367)	7861(378)	18(1)			
Female (number with clone)	4889(184)	3845(173)	8(0)			
Mean age (SD)	57(13)	59(10)				
Mean BMI (SD)	28(5.8)	29(5.6)				
Population-based						
JHS (Jackson Heart Study)						
Overall (number with clone)	1801(57)	589(25)	18(1)	Survival, malignancy, blood counts, cardiovascular events	The Jackson Heart Study is a population-based cohort of 5,301 African-Americans living in the Jackson, Mississippi metropolitan area. Subjects were enrolled as random members of community, volunteers, as part of the Atherosclerosis Risk in Communities (ARIC), or secondary family members. At 12-month intervals after the baseline clinic visit (Exam 1), participants were contacted by telephone to: update information; confirm vital statistics; document interim medical events, hospitalizations, and functional status; and obtain additional sociocultural information. Questions about medical events, symptoms of cardiovascular disease and functional status were repeated annually. Ongoing cohort surveillance includes abstraction of medical records and death certificates for relevant International Classification of Diseases (ICD) codes and adjudication of nonfatal events and deaths.	Wilson, J.G., et al. (2005). "Study design for genetic analysis in the Jackson Heart Study." <i>Ethnicity and Disease</i> 15:30-37; Taylor, H.A. (2005). "The Jackson Heart Study: An Overview." <i>Ethnicity and Disease</i> 15:1-3; Fuqua, S.R. et al., (2005). "Recruiting African-American Research Participation in the Jackson Heart Study: Methods, Response Rates, and Sample Description." <i>Ethnicity and Disease</i> 15:18-29; Fuqua, S.R. et al., (2005).
Female (number with clone)	1003(30)	401(18)	8(0)			
Mean age (SD)	51(13)	58(11)				
Mean BMI (SD)	31(6.8)	34(6.4)				
GoT2D						
Botnia (The Botnia Study), Diabetes_reg (Scania Diabetes Registry), Helsinki-sib (Helsinki siblings with diabetes cohort), Malmö-sib (siblings in Malmö, Sweden)						
Overall (number with clone)	224(12)	369(14)		Survival, blood counts, cardiovascular events	The Botnia Project was started in 1990 to study risk factors for T2D in western Finland and southern Sweden and includes ~11,000 people from ~1400 families. The Scania Diabetes Registry contains over 7000 diabetes patients recruited at hospitals in Scania, Sweden as from 1996. The majority of the patients come from the city of Malmö, and they account for about 25% of all diabetic patients in the region. Death information was obtained from the Finnish or Swedish Death and Hospital Discharge Registers. Individuals were ranked according to a liability model that measured risk for T2D. Briefly, liability scores were computed as the difference between diabetes status and the predicted risk based on age, BMI and gender; extreme cases were selected to have the highest liability scores (with diabetes but with low predicted risk for diabetes), and extreme controls were selected to have the lowest liability scores (without diabetes but with high predicted risk for diabetes).	Flannick, J., et al. (2013). "Assessing the phenotypic effects in the general population of rare variants in genes for a dominant Mendelian form of diabetes." <i>Nat Genet</i> 45(11): 1380-1385; Groop, L., et al. (1996). "Metabolic consequences of a family history of NIDDM (the Botnia study): evidence for sex-specific parental effects." <i>Diabetes</i> 45(11): 1585-1593; Lindholm, E., et al. (2001). "Classifying diabetes according to the new WHO clinical stages." <i>Eur J Epidemiol</i> 17(11): 983-989.
Female (number with clone)	108(6)	211(5)				
Mean age (SD)	65(10)	57(10)				
Mean BMI (SD)	30(3.7)	25(2.7)				
FUSION (Finland-United States Investigation of NIDDM Genetics Study)						
Overall (number with clone)	474(27)	470(21)		Survival, blood counts, cardiovascular events	The Finland-United States Investigation of NIDDM Genetics (FUSION) study is a long-term effort to identify genetic variants that predispose to type 2 diabetes (T2D) or that impact the variability of T2D-related quantitative traits. Unrelated T2D cases were selected from FUSION affected-sibpair families and from stage 2 replication. NGT controls with higher age and BMI were prioritized and were frequency matched to cases by birth province.	Valle, T., et al. (1998). "Mapping genes for NIDDM. Design of the Finland-United States Investigation of NIDDM Genetics (FUSION) Study." <i>Diabetes Care</i> 21(6): 949-958; Scott, L. et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. <i>Science</i> 316(5829), 1341-1345 (2007)
Female (number with clone)	213(11)	201(10)				
Mean age (SD)	63(7.2)	58(8.0)				
Mean BMI (SD)	28(3.9)	31(5.5)				
KORA						
Overall (number with clone)	91(11)	97(6)			KORA is a regional research platform for population-based studies, subsequent follow-up studies and family studies, established in 1996. Cases and controls were all of German ancestry. Cases were identified by self report of T2D in a personal interview which was validated by a questionnaire mailed to the treating physician and/or by medical chart review. Controls were non-diabetic as defined by self-report.	Wichmann, H. E., et al. (2005). "KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes." <i>Gesundheitswesen</i> 67 Suppl 1: S26-30.
Female (number with clone)	58(9)	43(3)				
Mean age (SD)	70(5.6)	61(8.1)				
Mean BMI (SD)	35(3.5)	28(2.8)				
MPP (Malmo Preventive Project)						
Overall (number with clone)	222(14)	110(8)			The MPP was started in the early 1970's as a screening survey in the middle-aged population of Malmö, the third largest city of Sweden. Subjects born in Malmö and residents of the city were invited for a clinical examination, questionnaire and blood sampling. In all 22,444 men and 10,902 women participated during the period 1974-1992. Cases and controls were selected as for Botnia.	Berglund, G., et al. (2000). "Long-term outcome of the Malmo preventive project: mortality and cardiovascular morbidity." <i>J Intern Med</i> 247(1): 19-29.
Female (number with clone)	87(7)	51(5)				
Mean age (SD)	68(5.2)	47(6.1)				
Mean BMI (SD)	36(1.8)	23(1.4)				
STT (UKT2D Consortium, Controls)						
Overall (number with clone)	319(14)	*			Non-diabetic controls selected from the Twins UK Study. A twin pair was considered for selection if there was no recorded family history of diabetes, neither twin was ever recorded as impaired glucose tolerant, there were available quantitative trait and genetic data, and no evidence of admixture. From set of qualifying twin pairs, the best control twin was selected from each pair with the lowest ratio of fasting glucose level to BMI across all readings.	Moayyeri, A., et al. (2013). "The UK Adult Twin Registry (TwinsUK Resource)." <i>Twin Res Hum Genet</i> 16(1): 144-149.
Female (number with clone)	265(11)	*				
Mean age (SD)	61(10)	*				
Mean BMI (SD)	31(5.9)	*				
SIGMA						
MEC (Multiethnic cohort, Hispanics in Los Angeles)						
Overall (number with clone)	445(24)	496(27)		Survival, malignancy, cardiovascular events	The MEC consists of 215,251 men and women in Hawai'i and Los Angeles. Those self-identified as Latinos were used for sequencing in a prior study. Between 1993 and 1996, adults between 45 and 75 years old were enrolled. Potential cohort members were identified through Department of Motor Vehicles drivers' license file, voter registration files and Health Care Financing Administration data files. Between 1995 and 2004, blood specimens were collected from ~67,000 MEC participants. Controls were frequency matched to cases on sex, ethnicity and age at entry into the cohort (5-year age groups) and place of birth (U.S. vs. Mexico, South or Central America). Persons in the cohort who develop cancer are identified through Surveillance, Epidemiology, and End Results (SEER) Program registries that have been established by state statute in Hawai'i and California.	Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, et al. (2000) A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. <i>Am J Epidemiol</i> 151: 346-357.
Female (number with clone)	228(13)	263(14)				
Mean age (SD)	68(7.2)	68(7.3)				
Mean BMI (SD)	27(4.3)	30(5.7)				
MexB1 (UNAM/INCMNSZ Diabetes Study)						
Overall (number with clone)	543(7)	550(20)			Cases were recruited at the outpatient diabetes clinic of the Department of Endocrinology and Metabolism of the Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán (INCMNSZ). All Mexican-mestizo individuals were invited to participate in the study. Control subjects were recruited from a cohort of adults aged 45 years or older among government	SIGMA T2D Consortium, et al. (2014). "Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico." <i>Nature</i> 506(7486): 97-101.

	Female (number with clone)	336(6)	326(13)	subjects were recruited from a cohort of individuals aged 40 years or older among government employees, blue collar workers and subjects seeking for attention in medical units for any condition besides those considered as exclusion criteria (diabetes, coronary heart disease, stroke, transient ischemic attack, lower limb amputations, alcoholism (more than 10 servings of alcohol per week) or any disease that in opinion of the researcher may limit life expectancy to less than 2 years). Diagnosis of type 2 diabetes was done following the American Diabetes Association (ADA) criteria	2011
	Mean age (SD)	55(9.4)	55(13)		
	Mean BMI (SD)	28(3.8)	28(4.4)		
MexB2 (Diabetes in Mexico Study)	Overall (number with clone)	177(3)	393(11)	Individuals were recruited from two tertiary level institutions (IMSS and ISSSTE) located in Mexico City. Unrelated healthy subjects older than 45 years and with fasting glucose levels below 100 mg/dL were classified as controls. Unrelated individuals, older than 18 years, with either previous T2D diagnosis or fasting glucose levels above 125 mg/dL were included as T2D cases.	SIGMA T2D Consortium, et al. (2014). "Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico." Nature 506(7486): 97-101.
	Female (number with clone)	134(3)	275(9)		
	Mean age (SD)	56(9.9)	57(12)		
	Mean BMI (SD)	28(4.6)	29(5.5)		
MexB3 (Mexico City Diabetes Study)	Overall (number with clone)	550(23)	281(9)	The Mexico City Diabetes Study is a population based prospective investigation. All 35-64 years of age men and non-pregnant women residing in the study site (low income neighborhoods equivalent to 6 census tracts with a total population of 15,000 inhabitants) were interviewed and invited to participate in the study. There was a response rate of 67% for the initial exam. Diagnostic criteria for type 2 diabetes were recommended by the ADA.	SIGMA T2D Consortium, et al. (2014). "Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico." Nature 506(7486): 97-101.
	Female (number with clone)	334(14)	164(3)		
	Mean age (SD)	62(7.6)	64(7.5)		
	Mean BMI (SD)	29(4.8)	30(5.5)		
T2D-GENES					
		No T2D	T2D		
AW (Wake Forest School of Medicine Study)	Overall (number with clone)	531(18)	529(48)	Cases are self-reported diabetics with diabetic nephropathy, recruited from dialysis clinics with age of onset ≥ 25 . Controls were recruited from community and internal medicine clinics and had a current diagnosis of diabetes or renal disease.	Palmer, N. D. et al. A genome-wide association search for type 2 diabetes genes in African Americans. PLoS One 7:e29202. (2012)
	Female (number with clone)	300(9)	317(28)		
	Mean age (SD)	51(12)	64(9.1)		
	Mean BMI (SD)	30(7.0)	29(6.6)		
EK (Korea Association Research Project)	Overall (number with clone)	554(33)	521(10)	Cases selected for age of onset of T2D ≥ 40 years. Participants with early onset T2D and family history were prioritized. Controls were selected for not having current or past diabetes and no diabetic medications. Older subjects with normal glucose were prioritized.	Cho, Y. S. et al. A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. Nat. Genet. 41, 527-534 (2009)
	Female (number with clone)	324(18)	237(4)		
	Mean age (SD)	63(3.6)	54(7.5)		
	Mean BMI (SD)	24(3.1)	26(3.3)		
Study and Singapore Prospective Study Program	Overall (number with clone)	592(25)	476(24)	Cases were clinically ascertained T2D from primary care clinics. Individuals with early age of diagnosis and with at least one first degree relative with T2D were preferentially selected. Controls were defined as: fasting blood glucose < 6 mmol/L, no personal history of diabetes, and no anti-diabetic medication. Older controls preferentially selected.	Sim, X. et al. Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. PLoS Genet. 7(4), e1001363 (2011)
	Female (number with clone)	363(11)	250(11)		
	Mean age (SD)	58(7.0)	58(9.3)		
	Mean BMI (SD)	23(3.4)	26(3.8)		
HA (Hispanics from San Antonio Family Heart Study, San Antonio Family Diabetes/Gallbladder Study, Veterans Administration Genetic Epidemiology Study, and the Investigation of Nephropathy and Diabetes Study family component)	Overall (number with clone)	111(6)	142(3)	Cases were drawn from four separate family studies and met the following criteria: ADA 2002 criterion, WHO 1999 criteria, or physician reported diagnosis with current medical therapy. Controls defined by not having fasting glucose < 126 mg/dL at each visit and no history of prior diabetic medication.	Mitchell, B. D. et al. Genetic and environmental contributions to cardiovascular risk factors in Mexican Americans. The San Antonio Family Heart Study. Circulation 94, 2159-2170 (1996); Hunt, K. J. et al. Genome-wide linkage analyses of type 2 diabetes in Mexican Americans: the San Antonio Family Diabetes/Gallbladder Study. Diabetes 54, 2655-2662 (2005); Coletta, D. K. et al. Genome-wide linkage scan for genes influencing plasma triglyceride levels in the Veterans Administration Genetic Epidemiology Study. Diabetes 58, 279-284 (2009); Knowler, W. C. et al. The Family Investigation of Nephropathy and Diabetes (FIND): design and methods. J. Diabetes Complicat. 19, 1-9 (2005)
	Female (number with clone)	62(4)	90(1)		
	Mean age (SD)	44(14)	51(12)		
	Mean BMI (SD)	30(5.5)	33(6.3)		
HS (Hispanics in Starr County, Texas)	Overall (number with clone)	704(3)	751(32)	Cases defined by fasting glucose ≥ 140 mg/dl on more than 1 occasion or self-reported physician-diagnosed diabetes with current medical therapy. In instances where cases were drawn from families, the individual with youngest age at onset was chosen. Controls ascertained from epidemiologically represented sample of individuals in Starr County, TX with individuals with known diagnosis of diabetes excluded. Controls are significantly younger on average.	Hanis, C. L. et al. Diabetes among Mexican Americans in Starr County, Texas. Am. J. Epidemiol. 118, 659-672 (1983); Below JE, et al. Genome-wide association and meta-analysis in populations from Starr County, Texas and Mexico City identify type 2 diabetes susceptibility loci and enrichment for eQTLs in top signals. Diabetologia 54, 2047-2055 (2011)
	Female (number with clone)	506(3)	449(15)		
	Mean age (SD)	39(9.9)	56(12)		
	Mean BMI (SD)	30(6.2)	32(6.4)		
SL (London Life Sciences Population Study [UK Indian Asians])	Overall (number with clone)	538(24)	531(16)	A population-based cohort study of Indian Asians living in West London, UK with all 4 grandparents born on the Indian subcontinent. Prevalent T2D defined as previous physician diagnosis of diabetes on treatment, with onset of diabetes after the age of 18 years and without insulin use in the first year after diagnosis; or fasting plasma glucose > 7.0 mmol/L. Controls defined as no previous history of diabetes, no anti-diabetic medication, and fasting plasma glucose < 6.0 mmol/L.	Chambers, J.C. et al. Genome-wide association study identifies variants in TMPRSS6 associated with hemoglobin levels. Nat. Genet. 41, 1170-1172 (2009); Chambers, J.C. et al. Common genetic variation near melatonin receptor MTRNR1B contributes to raised plasma glucose and increased risk of type 2 diabetes among Indian Asians and European Caucasians. Diabetes 58, 2703-2708 (2009); van der Harst, P. et al. Seventy-five genetic loci influencing the human red blood cell. Nature 492, 369-375 (2012)
	Female (number with clone)	85(4)	75(2)		
	Mean age (SD)	63(9.2)	53(5.6)		
	Mean BMI (SD)	27(3.5)	27(2.9)		
SS (Singapore Indian Eye Study [Singapore Indians])	Overall (number with clone)	585(15)	563(31)	Cases selected for HbA1c $> 6.5\%$ or personal history of diabetes with age at diagnosis available. Preferentially selected cases with at least one first degree relative with T2D. Controls selected for HbA1c $< 6\%$, no personal history of diabetes, and not taking antidiabetic medication. Older controls preferentially selected.	Sim, X. et al. Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. PLoS Genet. 7(4), e1001363 (2011)
	Female (number with clone)	288(5)	250(14)		
	Mean age (SD)	56(10)	61(9.7)		
	Mean BMI (SD)	25(4.8)	27(5.1)		

UA (Ashkenazis)			Survival, blood counts	Subjects in this cohort are of Ashkenazi Jewish origin, defined as having all four grandparents born in Northern or Eastern Europe; subjects with known or suspected Sephardic Jewish or non-Jewish ancestry excluded. T2D cases were selected from two separate DNA collections: 1. Genome-wide, affected-sibling-pair linkage study (Permutt et al. Diabetes 2001) or 2. Study to determine genetic risk for diabetic complications (Blech et al. PLoS One 2011). Controls were selected for fasting blood glucose <7 mmol/L, no personal history of diabetes, and no anti-diabetic medications. Controls included elderly (>90 years) individuals who were part of the Longevity Genes Project.	Atzmon, G. et al. Lipoprotein genotype and conserved pathway for exceptional longevity in humans. PLoS Biol. 4(4), e113 (2006); Atzmon, G. et al. Evolution in health and medicine Sackler colloquium: Genetic variation in human telomerase is associated with telomere length in Ashkenazi centenarians. Proc Natl Acad Sci U S A. 107 (Suppl 1), 1710- 1717 (2010); Permutt, M.A. et al. A genome scan for type 2 diabetes susceptibility loci in a genetically isolated population. Diabetes 50(3), 681-685 (2001); Blech et al. Predicting diabetic nephropathy using a multifactorial genetic model. PLoS One 6(4), e18743 (2011)
Overall (number with clone)	342(42)	506(49)			
Female (number with clone)	195(20)	242(18)			
Mean age (SD)	79(13)	66(8.6)			
Mean BMI (SD)	25(4.3)	27(3.2)			
UM (Metabolic Syndrome in Men Study)			The METSIM Study includes 10,197 men, aged from 45 to 73 years, randomly selected from the population register of the town of Kuopio, Eastern Finland, and examined in 2005-2010. The aim of the study is to investigate genetic and non-genetic factors associated with the risk of type 2 diabetes (T2D), cardiovascular disease (CVD), and insulin resistance-related traits in a cross-sectional and longitudinal setting. Unrelated T2D cases with family history of diabetes were selected. Unrelated NGT controls were selected, prioritizing older individuals with no family history of diabetes.	Stancakova, A. et al. Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men. Diabetes 58, 1212-1221 (2009)	
Overall (number with clone)	500(9)	487(24)			
Female (number with clone)	*	*			
Mean age (SD)	55(4.6)	60(6.7)			
Mean BMI (SD)	26(3.2)	31(5.1)			

Supplementary Table S2
List of hematopoietic genes and variants queried

Gene name	Reported mutations used for variant calling	Accession	Number of variants found
ARID1A	Frameshift/nonsense, A305V, M872T	NM_006015	0
ASXL1	Frameshift/nonsense in exon 11-12	NM_015338	62
BCL10	Frameshift/nonsense/splice-site	NM_003921	0
BCL11B	Frameshift/nonsense/splice-site, A360T, C432Y, H445Y, R447H, G452K, H479Y, G596S, L617Q, G847R	NM_138576	1
BCL6	R40H, Y111H, P341H, S350R, R585W, Q679K	NM_001130845	0
BCOR	Frameshift/nonsense/splice-site	NM_001123385	4
BCORL1	Frameshift/nonsense/splice-site	NM_021946	2
BIRC3	Frameshift/nonsense/splice-site exon 2	NM_182962	1
BRAF	G464E, G464V, G466E, G466V, G469R, G469E, G469A, G469V, V471F, V472S, N581S, I582M, I592M, I592V, D594N, D594G, D594V, D594E, F595L, F595S, G596R, L597V, L597S, L597Q, L597R, A598V, V600M, V600L, V600K, V600R, V600E, V600A, V600G, V600D, K601E, K601N, R603*, W604R, W604G, S605G, S605F, S605N, G606E, G606A, G606V, H608R, H608L, G615R, S616P, S616F, L618S, L618W	NM_004333	2
BRCC3	Frameshift/nonsense/splice-site	NM_024332	6
BTG1	M1I, H2Y, P3R, Y5H, M11I, S23A, F25C, R27H, K29Q, L31F, Q36H, L37M, Q38E, F40C, E46D, E46Q, A49P, P58L, E59D, L94V, I115V, E117D, N165S	NM_001731	0
BTG2	A45T, A45E	NM_006763	0
CARD11	E93D, G123S, G126D, T128M, F130I, R179W, K182N, M183L, K215M, D230N, L232LI, M240MGLNKM, K244T, S250P, S250F, L251P, L251F, V266*, D338G, T353*, D357V, Y361H, M365K, D387V, D387T, D401V, R418S, R423W, E432K, E626K	NM_032415	1
CBL	RING finger missense p.381-421	NM_005188	12
CBLB	RING finger missense p.372-412	NM_170662	0
CCND3	Frameshift/nonsense/splice-site, T211A, P212L, V215G	NM_001760	0
CD58	Frameshift/nonsense/splice-site, G210C, G210S	NM_001779	1
CD70	L60R, G66R, F186S	NM_001252	0
CD79A	del191-226, del179-226, del191-208	NM_001783	0
CD79B	V9A, D89G, Y92F, D193G, D194G, Y196C, Y197C, Y197D, Y197H, T207fs, Y208*, V212A, del195_197, del196_229, del193_229, del205_229	NM_000626	1
CDKN2A	Frameshift/nonsense/splice-site	NM_000077	0
CDKN2B	Frameshift/nonsense/splice-site	NM_004936	0
CEBPA	Frameshift/nonsense/splice-site	NM_004364	0
CHD2	H620L, F1146L, L1270F	NM_001271	0
CNOT3	Frameshift/nonsense, E20K, R57W, R57Q, E70K	NM_014516	1
CREBBP	Frameshift/nonsense/splice-site, D1435E, R1446L, R1446H, R1446C, Y1450C, P1476R, Y1482H, H1487Y, W1502C, Y1503D, Y1503H, Y1503F, S1680del	NM_004380	6
CRLF2	F232C	NM_022148	0
CSF1R	L301F, L301S, Y969C, Y969N, Y969F, Y969H, Y969D	NM_005211	0
CSF3R	T615A, T618I, truncating c.741-791	NM_000760	0
CTCF	Frameshift/nonsense, R377C, R377H, P378A, P378L	NM_006565	0
CUX1	Frameshift/nonsense	NM_181552	3
DDX3X	R276K, R376S, R376C, D506V, R528C, R528H, R534S, R534H, P568L	NM_001356	2
DIS3	R780K, R780T, R780H, R780S, D488H, P480Q, D488N, S477R, R467Q, M662R, R689Q, R514K	NM_014953	0
DNMT3A	Frameshift/nonsense/splice-site, P307S, P307R, R326H, R326L, R326C, R326S, R366P, R366H, R366G, A368T, F414L, F414S, F414C, C497Y, Q527H, Q527P, Y533C, G543A, G543S, G543C, L547H, L547P, L547F, M548I, M548K, G550R, W581R, W581G, W581C, G646V, G646E, L653W, L653F, V657A, V657M, R659H, Y660C, R676W, R676Q, G685R, G685E, G685A, D686Y, D686G, G699R, G699S, G699D, P700S, P700R, P700Q, D702N, D702Y, V704M, V704G, I705F, I705T, I705S, C710S, S714C, N717S, N717I, P718L, R720H, R720G, Y724C, R729Q, R729W, R729G, F731L, F732del, F732S, F732L, F734L, F734C, Y735C, Y735N, Y735S, R736H, R736C, R736P, L737H, L737V, L737F, L737R, A741V, R749C, R749L, F751L, F752del, F752C, F752L, F752I, F752V, L754R, L754H, F755S, F755I, F755L, M761I, M761V, G762C, S770W, S770P, R771Q, F772I, F772V, L773R, E774K, E774D, D781G, R792H, G796D, G796V, N797Y, N797H, P799R, P799H, R803S, P804S, P804L, S828N, K829R, Q842E, P849L, D857N, W860R, F868S, G869S, G869V, M880V, S881R, S881I, R882H, R882P, R882C, R882G, Q886R, G890D, L901R, L901H, P904L, F909C, A910P	NM_022552	403
EBF1	Frameshift/nonsense, M1R, S7G, F54S, G143V, G171D, Q196P, N237K, S238T, S238Y, R381S	NM_024007	0
EED	Frameshift/nonsense/splice-site, L240Q, I363M	NM_003797	0
EP300	Frameshift/nonsense/splice-site, VF1148_1149del, D1399N, D1399Y, P1452L, Y1467N, Y1467H, Y1467C, R1627W, A1629V	NM_001429	1
ETV6	Frameshift/nonsense/splice-site	NM_001987	1
EZH2	Frameshift/nonsense/splice-site, Q62R, N102S, F145S, F145C, F145Y, F145L, G159R, E164D, R202Q, K238E, E244K, R283Q, H292R, P488S, R497Q, R561H, T568I, K629E, Y641N, Y641H, Y641S, Y641C, Y641F, D659Y, D659G, V674M, A677G, A677V, R679C, R679H, R685C, R685H, A687V, N688I, N688K, H689Y, S690P, I708V, I708T, I708M, E720K, E740K	NM_001203247	2
EZR	I245M	NM_003379	0
FAM46C	Frameshift/nonsense/splice-site	NM_017709	1
FAS	Frameshift/nonsense/splice-site	NM_000043	0
FBXO11	Frameshift/nonsense/splice-site	NM_001190274	0
FBXW7	Frameshift/nonsense/splice-site, E74A, D101V, F280L, R465H, R505C, G597E, R1165Q	NM_033632	1
FLT3	V579A, V592A, V592I, F594L, M737I, FY590-591GD	NM_004119	1
FOXP1	Frameshift/nonsense/splice-site	NM_032682	1
FYN	L174R, R176C, Y531H	NM_002037	0
GATA1	Frameshift/nonsense/splice-site	NM_002049	0
GATA2	Frameshift/nonsense/splice-site, R293Q, N317H, A318T, A318V, A318G, G320D, L321P, L321F, L321V, Q328R, R330Q, R361L, L359V, A372T, R384G, R384K	NM_001145661	0
GATA3	Frameshift/nonsense/splice-site ZNF domain, R276W, R276Q, N286T, L348V,	NM_001002295	0
GNA13	I34T, G57S, S62F, M68K, Q134R, Y145F, L152F, E167D, Q169H, R264H, E273K, V322G, V362G, L371F	NM_006572	0
GNAS	R201(844)S, R201(844)C, R201(844)H, R201(844)L, Q227(870)K, Q227(870)R, Q227(870)L, Q227(870)H, R374(1017)C	NM_016592	8
GNB1	K57N, K57M, K57E, K57T, I80T, I80N	NM_002074	22
HIST1H1B	S89N, S89R, G101D, G73A, K84N, A123D	NM_005322	0
HIST1H1C	P118S, P129A, K156R, K187R, K/G81/83N/A	NM_005319	1

HIST1H1D	Frameshift/nonsense	NM_005320	0
HIST1H1E	A158T, A167V, P196S, K202E, K205R	NM_005321	0
HIST1H3B	A48S, S87N, S87T	NM_003537	0
HLA-A	Frameshift/nonsense, G124D, A164T	NM_002116	0
ID3	Frameshift/nonsense/splice-site, S39R, V55Q, P56S, L64F, S65R, I74V, L80R, H96R,	NM_002167	0
IDH1	R132C, R132G, R132H, R132L, R132P, R132V, V178I	NM_005896	0
IDH2	R140W, R140Q, R140L, R140G, R172W, R172G, R172K, R172T, R172M, R172N, R172S	NM_002168	3
IKBKB	K171E	NM_001556	0
IKZF1	Frameshift/nonsense	NM_006600	1
IKZF2	Frameshift/nonsense	NM_016260	0
IKZF3	Frameshift/nonsense	NM_012481	0
IL7R	exon 6 cysteine insertion	NM_002185	0
INTS12	M445I	NM_020395	0
IRF4	N2S, S18T, I32V, L40V, Q60K, Q60H	NM_002460	0
IRF8	Frameshift/nonsense from c.377-426, S34T, S55A, T80A, K108E,	NM_002163	0
JAK1	T478A, T478S, V623A, A634D, L653F, R724H, R724Q, T782M, L783F	NM_002227	0
JAK2	N533D, N533Y, N533S, H538R, K539E, K539L, I540T, I540V, V617F, R683S, R683G, del/ins537-539L, del/ins538-539L, del/ins540-543MK, del/ins540-544MK, del/ins541-543K, del542-543, del543-544, ins11546-547	NM_004972	31
JAK3	M511T, M511I, A572V, A572T, A573V, R657Q, V715I, V715A	NM_000215	1
JARID2	Frameshift/nonsense/splice-site	NM_004973	1
KDM6A	Frameshift/nonsense/splice-site, del419	NM_021140	2
KIT	ins503, V559A, V559D, V559G, V559I, V560D, V560A, V560G, V560E, del560, E561K, del579, P627L, P627T, R634W, K642E, K642Q, V654A, V654E, H697Y, H697D, E761D, K807R, D816H, D816Y, D816F, D816L, D816V, D816H, del551-559	NM_000222	1
KLHL6	F49L, L58P, T64S, T64I, L65V, L65P, Q81, L90V, R98T, R98W	NM_130446	1
KRAS	G12D, G12A, G12E, G12V, G13D, G13C, G13Y, G13F, G13R, G13A, G13V, G13E, T58I, G60D, G60A, G60V, Q61K, Q61E, Q61P, Q61R, Q61L, Q61H, K117E, K117N, A146T, A146P, A146V	NM_033360	3
LEF1	Frameshift/nonsense	NM_016269	0
LRRK2	E155K, I543S	NM_198578	0
LTB	Frameshift/nonsense	NM_002341	0
LUC7L2	Frameshift/nonsense/splice-site	NM_016019	3
MALT1	V381F, Y750S	NM_006785	0
MAP2K1	F53L, Q56P, K57T, K57N, K57E, I103N, C121S, N122D, P124Q	NM_002755	0
MAP3K14	H683Q	NM_003954	0
MED12	E33K, L36P, G44S, A59P, R521H, L1224F	NM_005120	0
MEF2B	Frameshift/nonsense/splice-site, I8F, L67R, Y69C, Y69H, E77K, N81K, N81Y, D83V, D83A	NM_001145785	0
MLL	Frameshift/nonsense	NM_005933	0
MLL2	Frameshift/nonsense	NM_003482	4
MPL	S505G, S505N, S505C, L510P, del513, W515A, W515R, W515K, W515S, W515L, A519T, A519V, Y591D, W515-518KT	NM_005373	2
MXRA5	Frameshift/nonsense/splice-site	NM_015419	0
MYD88	V217F, S219C, M240T, S251N, P266, L273P	NM_001172567	2
NOTCH1	Frameshift/nonsense	NM_017617	2
NOTCH2	Frameshift/nonsense	NM_024408	2
NPM1	Frameshift p.W288fs (insertion at c.859_860,860_861,862_863,863_864)	NM_002520	0
NRAS	G12S, G12R, G12C, G12N, G12P, G12Y, G12D, G12A, G12V, G12E, G13S, G13R, G13C, G13N, G13P, G13Y, G13D, G13A, G13V, G13E, G60E, G60R, Q61R, Q61L, Q61K, Q61P, Q61H, Q61Q	NM_002524	6
P2RY8	M52K, Y82C, R86C, C144Y, M189I, A251T, N254S, F255I, V256M	NM_178129	0
PAPD5	Frameshift/nonsense	NM_001040284	0
PAX5	Frameshift/nonsense/splice-site, S66N, P80R,	NM_016734	0
PDS5B	Frameshift/nonsense/splice-site, R1292Q	NM_015032	0
PDSS2	Frameshift/nonsense	NM_020381	1
PHF6	Frameshift/nonsense/splice-site, A40D, M125I, S246Y, F263L, R274Q, C297Y, H302Y, H329L	NM_001015877	1
PIK3CA	I543V, Q545G, G1007D, L1026P, M1040I, D1045N, H1047R	NM_006218	1
POT1	Frameshift/nonsense/splice-site before OB domain (c.274), M1L, Y36N, K90E, Q94R, Y223C, H266L, G272V, C591W	NM_015450	0
POU2AF1	P27A	NM_006235	0
POU2F2	T223A, T223S, T239S, R282H, T307I, G392R	NM_002698	0
PRDM1	Frameshift/nonsense/splice-site	NM_001198	3
PRPF40B	Frameshift/nonsense/splice-site, P15H, M58I, P405L, P562S,	NM_001031698	1
PRPF8	M1307I	NM_006445	0
PTEN	Frameshift/nonsense/splice-site, D24G, R47G, F56V, L57W, H61R, K66N, Y68H, C71Y, F81C, Y88C, D92G, D92V, D92E, H93Y, H93D, H93Q, N94I, P95L, I101T, C105F, C105S, D107Y, L112V, H123Y, C124R, C124S, K125E, A126D, K128N, R130G, R130Q, R130L, G132D, I135V, I135K, C136R, C136F, K144Q, A151T, D153Y, D153N, Y155H, Y155C, R159K, R159S, R161K, R161I, G165R, G165E, S170N, S170I, R173C, Y174D, Y177C, H196Y, R234W, G251C, D252Y, F271S, D326G	NM_000314	0
PTPN1	exon 1 frameshift/nonsense, Q9E	NM_002827	0
PTPN11	G60V, G60R, G60A, D61Y, D61V, D61G, Y63C, E69K, E69G, E69D, E69Q, F71L, F71K, A72T, A72V, A72D, T73I, E76K, E76Q, E76M, E76A, E76G, E139G, E139D, N308D, N308T, N339S, P491L, S502P, S502A, S502L, G503V, G503G, G503A, G503E, Q506P, T507A, T507K	NM_002834	3
RAD21	Frameshift/nonsense/splice-site, R65Q, H208R, Q474R	NM_006265	6
RBBP4	E330K	NM_005610	0
RHOA	C16R, G17V, G17E, T19I, D120Y	NM_001664	0
RIT1	Frameshift/nonsense, Q79, E81Q, E81G, F82L, F82C, F82I, F82V, M90I, R122L	NM_006912	1
RPL10	R98S, Q123P	NM_006013	0
RPL5	Frameshift/nonsense/splice-site, Q63R	NM_000969	0
RPS15	G132S, A135V, A135G, S139A, K145N	NM_001018	1
RPS2	R200G	NM_002952	0
RUNX1	Frameshift/nonsense/splice-site, S73F, H78Q, H78L, R80C, R80P, R80H, L85Q, P86L, P86H, S114L, D133Y, L134P, R135G, R135K, R135S, R139Q, R142S, A165V, R174Q, R177L, R177Q, A224T, D171G, D171V, D171N, R205W, R293Q	NM_001001890	0

SF3A1	Frameshift/nonsense/splice-site, A57S, M117I, K166T, Y271C	NM_005877	1
SF3B1	G347V, R387W, R387Q, E592K, E622D, Y623C, R625L, R625C, H662Q, H662D, K666N, K666T, K666E, K666R, K700E, V701F, A708T, G740R, G740E, A744P, D781G, E783K	NM_012433	27
SFRS2	Y44H, P95H, P95L, P95T, P95R, P95A, P107H, P95fs	NM_003016	11
SGK1	Frameshift/nonsense/splice-site	NM_001143676	0
SMC1A	K190T, R586W, M689V, R807H, R1090H, R1090C	NM_006306	1
SMC3	Frameshift/nonsense, R155I, Q367E, D392V, K571R, R661P, G662C	NM_005445	1
SOCS1	Frameshift/nonsense/splice-site, R48W	NM_003745	0
SPRY4	I113N, G117R	NM_001127496	0
STAG1	Frameshift/nonsense/splice-site, H1085Y	NM_005862	1
STAG2	Frameshift/nonsense/splice-site	NM_006603	1
STAT3	M206K, G618R, Y640F, N642H, N647I, D661N, D661H, D661Y, D661V	NM_139276	4
STAT5A	N642H	NM_003152	0
STAT5B	N642H, Y665F	NM_012448	0
STAT6	N417Y, N417S, D419H, D419G, D419V, N421K, N430T, N430S	NM_001178081	0
SUZ12	Frameshift/nonsense	NM_015355	1
SWAP70	Frameshift/nonsense/splice-site	NM_015055	0
TBL1XR1	Frameshift/nonsense/splice-site	NM_024665	1
TCF3	N551K, V557E, V557G, D561E, D561V, D561N, M572K	NM_003200	0
TET1	Frameshift/nonsense/splice-site, V128F, H1297Y, R1656C, V2120M,	NM_030625	1
TET2	Frameshift/nonsense/splice-site, S282F, N312S, L346P, S460F, D666G, P941S, C1135Y	NM_001127208	72
TMEM30A	Frameshift/nonsense	NM_018247	0
TNF	L47F, H52Y, P58S	NM_000594	0
TNFAIP3	Frameshift/nonsense, D117V, M476I, P574L	NM_006290	3
TNFRSF14	Frameshift/nonsense/splice-site, Y26N, Y26D, C42P, C42W, A102P	NM_003820	2
TP53	Frameshift/nonsense/splice-site, S46F, G105C, G105R, G105D, G108S, G108C, R110L, R110C, T118A, T118R, T118I, S127F, S127Y, L130V, L130F, K132Q, K132E, K132W, K132R, K132M, K132N, C135W, C135S, C135F, C135G, C135Y, Q136K, Q136E, Q136P, Q136R, Q136L, Q136I, Q136H, A138P, A138V, A138A, A138T, T140I, C141R, C141G, C141A, C141Y, C141S, C141F, C141W, V143M, V143A, V143E, L145Q, W146G, W146L, L145R, V147G, P151T, P151A, P151S, P151H, P151R, P152S, P152R, P152L, T155P, V157F, R158H, R158L, A159V, A159P, A159S, A159D, A161T, A161D, Y163N, Y163H, Y163D, Y163S, Y163C, K164E, K164M, K164N, K164P, H168Y, H168P, H168R, H168L, H168Q, M169I, M169T, M169V, E171K, E171Q, E171G, E171A, E171V, E171D, V172D, V173M, V173L, V173G, R174W, R175G, R175C, R175H, C176R, C176G, C176Y, C176F, C176S, P177R, P177L, P177H, H178D, H178P, H178Q, H179Y, H179R, H179Q, R181C, R181Y, D186G, G187S, P190L, P190T, H193N, H193P, H193L, H193R, L194F, L194R, I195F, I195N, I195T, V197L, G199V, Y205N, Y205C, Y205H, D208V, R213Q, R213P, R213L, R213Q, H214D, H214R, S215G, S215I, S215R, V216M, V217G, Y220N, Y220H, Y220S, Y220C, E224D, I232F, I232N, I232T, I232S, Y234N, Y234H, Y234S, Y234C, Y236N, Y236H, Y236C, M237V, M237K, M237I, C238R, C238G, C238Y, C238W, N239T, N239S, S241Y, S241C, S241F, C242G, C242Y, C242S, C242F, G244S, G244C, G244D, G245S, G245R, G245C, G245D, G245A, G245V, G245S, M246V, M246K, M246R, M246I, N247I, R248W, R248G, R248Q, R249G, R249W, R249T, R249M, P250L, I251N, L252P, I254S, I255F, I255N, I255S, L257Q, L257P, E258K, E258Q, D259Y, S261T, G262D, G262V, L265P, G266R, G266E, G266V, R267W, R267Q, R267P, E271K, V272M, V272L, R273S, R273G, R273C, R273H, R273P, R273L, V274F, V274D, V274A, V274G, V274L, C275Y, C275S, C275F, A276P, C277F, P278T, P278A, P278S, P278H, P278R, P278L, G279E, R280G, R280K, R280T, R280I, R280S, D281N, D281H, D281Y, D281G, D281E, R282G, R282W, R282Q, R282P, E285K, E285V, E286G, E286V, E286K, K320N, L330R, G334V, R337C, R337L, A347T, L348F, T377P	NM_001126112	33
TRAF3	Frameshift/nonsense	NM_145725	0
TYW1	R555Q, E601K	NM_018264	0
U2AF1	D14G, S34F, S34Y, R35L, R156H, R156Q, Q157R, Q157P	NM_006758	5
U2AF2	R18W, Q143L, M144I, L187V, Q190L	NM_007279	0
UBR5	Frameshift/nonsense/splice-site exon 58	NM_015902	0
WT1	Frameshift/nonsense/splice-site	NM_024426	0
XBP1	L167I, P326R	NM_001079539	0
XPO1	E571A, E571K	NM_003400	0
ZNF471	G483, P486, R423	NM_020813	0
ZRSR2	Frameshift/nonsense, R126P, E133G, C181F, H191Y, I202N, F239V, F239Y, N261Y, C280R, C302R, C326R, H330R, N382K	NM_005089	3
Total			805

Supplementary Table S3
Called somatic variants in 160 hematopoietic genes

Gene name	Chrom	Start position	Variant Type	Reference Allele	Variant Allele	Variant Classification	Protein Change	cDNA change	Accession	Variant allele fraction	Variant allele count	Reference allele count	Reference ID
ASXL1	20	31021211	SNP	C	T	Nonsense Mutation	p.R404*	c.1210C>T	NM_015338	0.302326	26	60	id16093
ASXL1	20	31021211	SNP	C	T	Nonsense Mutation	p.R404*	c.1210C>T	NM_015338	0.106796	11	92	id5983
ASXL1	20	31021211	SNP	C	T	Nonsense Mutation	p.R404*	c.1210C>T	NM_015338	0.063291	5	74	id3697
ASXL1	20	31021211	SNP	C	T	Nonsense Mutation	p.R404*	c.1210C>T	NM_015338	0.194444	14	58	id6530
ASXL1	20	31021211	SNP	C	T	Nonsense Mutation	p.R404*	c.1210C>T	NM_015338	0.052083	5	91	id12539
ASXL1	20	31021250	SNP	C	T	Nonsense Mutation	p.R417*	c.1250G>A	NM_015338	0.04878	6	117	id6912
ASXL1	20	31021250	SNP	C	T	Nonsense Mutation	p.R417*	c.1250G>A	NM_015338	0.052632	5	90	id17093
ASXL1	20	31021250	SNP	C	T	Nonsense Mutation	p.R417*	c.1250G>A	NM_015338	0.241379	28	88	id7330
ASXL1	20	31021250	SNP	C	T	Nonsense Mutation	p.R417*	c.1250G>A	NM_015338	0.177305	25	116	id4693
ASXL1	20	31021250	SNP	C	T	Nonsense Mutation	p.R417*	c.1250G>A	NM_015338	0.069767	6	80	id10535
ASXL1	20	31021276	DEL	C	-	Frame Shift Del	p.Y425fs	c.1275delC	NM_015338	0.2	16	65	id17167
ASXL1	20	31021283	SNP	C	T	Nonsense Mutation	p.Q428*	c.1282C>T	NM_015338	0.051948	4	73	id15889
ASXL1	20	31021295	SNP	C	T	Nonsense Mutation	p.Q432*	c.1294C>T	NM_015338	0.232877	17	56	id12542
ASXL1	20	31021332	SNP	C	G	Nonsense Mutation	p.S444*	c.1331C>T	NM_015338	0.041237	4	93	id9683
ASXL1	20	31021332	SNP	C	G	Nonsense Mutation	p.S444*	c.1331C>T	NM_015338	0.047619	7	140	id3121
ASXL1	20	31021535	SNP	C	T	Nonsense Mutation	p.Q512*	c.1534C>T	NM_015338	0.122449	24	172	id3907
ASXL1	20	31021538	SNP	G	T	Nonsense Mutation	p.E513*	c.1537G>T	NM_015338	0.052632	7	126	id12605
ASXL1	20	31021543	DEL	TG	-	Frame Shift Del	p.T514fs	c.1542_1543delTG	NM_015338	0.24	58	179	id9964
ASXL1	20	31021543	DEL	TG	-	Frame Shift Del	p.T514fs	c.1542_1543delTG	NM_015338	0.09	20	201	id14785
ASXL1	20	31021550	SNP	C	T	Nonsense Mutation	p.Q517*	c.1549C>T	NM_015338	0.067873	15	206	id15925
ASXL1	20	31021553	SNP	G	T	Nonsense Mutation	p.E518*	c.1552_1553delGA	NM_015338	0.095238	12	114	id15929
ASXL1	20	31021553	DEL	GA	-	Frame Shift Del	p.E518fs	c.1552_1553delGA	NM_015338	0.16	41	220	id16839
ASXL1	20	31021586	SNP	C	T	Nonsense Mutation	p.Q529*	c.1585C>T	NM_015338	0.179775	32	146	id15189
ASXL1	20	31021627	INS	-	GAAGATC	Frame Shift Ins	p.L542fs	AGATC	NM_015338	0.13	14	91	id16126
ASXL1	20	31021637	SNP	C	T	Nonsense Mutation	p.Q546*	c.1636C>T	NM_015338	0.194444	21	87	id15951
ASXL1	20	31021647	INS	-	TAACACT	Frame Shift Ins	p.R549fs	ACACT	NM_015338	0.05	10	206	id2117
ASXL1	20	31022238	SNP	C	T	Nonsense Mutation	p.Q575*	c.1723C>T	NM_015338	0.166667	10	50	id2349
ASXL1	20	31022263	SNP	G	A	Nonsense Mutation	p.W583*	c.1748G>A	NM_015338	0.08	6	69	id11045
ASXL1	20	31022286	INS	-	A	Frame Shift Ins	p.Y591fs	c.1771_1772insA	NM_015338	0.32	29	61	id5794
ASXL1	20	31022288	SNP	C	A	Nonsense Mutation	p.Y591*	c.1771_1772insA	NM_015338	0.101695	18	159	id13763
ASXL1	20	31022297	SNP	C	A	Nonsense Mutation	p.C594*	c.1782C>A	NM_015338	0.16129	10	52	id884
ASXL1	20	31022363	INS	-	T	Frame Shift Ins	p.D616fs	c.1848_1849insT	NM_015338	0.23	10	33	id1924
ASXL1	20	31022403	DEL	CACCACTGC CATAGAGA GGCGGC	-	Frame Shift Del	p.H630fs	c.1888_1910delCA CCACTGCATAGAG AGCGCGC	NM_015338	0.18	8	37	id6015
ASXL1	20	31022416	DEL	GAGAGGCG GCCCACT GCCATC	-	Frame Shift Del	p.R634fs	c.1901_1923delGA GAGGCGGCCACCA	NM_015338	0.22	13	45	id4773
ASXL1	20	31022441	INS	-	G	Frame Shift Ins	p.G642fs	c.1926_1927insG	NM_015340	0.33	6	12	id13857
ASXL1	20	31022441	INS	-	G	Frame Shift Ins	p.G642fs	c.1926_1927insG	NM_015338	0.24	6	19	id15227
ASXL1	20	31022441	INS	-	G	Frame Shift Ins	p.G642fs	c.1926_1927insG	NM_015339	0.38	6	10	id3122
ASXL1	20	31022441	INS	-	G	Frame Shift Ins	p.G642fs	c.1926_1927insG	NM_015341	0.38	8	13	id10828
ASXL1	20	31022441	INS	-	G	Frame Shift Ins	p.G642fs	c.1926_1927insG	NM_015342	0.31	11	24	id13059
ASXL1	20	31022589	SNP	C	T	Nonsense Mutation	p.Q692*	c.2074C>T	NM_015338	0.067416	6	83	id12078
ASXL1	20	31022592	SNP	C	T	Nonsense Mutation	p.R693*	c.2077C>T	NM_015338	0.108108	4	33	id12402
ASXL1	20	31022592	SNP	C	T	Nonsense Mutation	p.R693*	c.2077C>T	NM_015338	0.166667	5	25	id5859
ASXL1	20	31022592	SNP	C	T	Nonsense Mutation	p.R693*	c.2077C>T	NM_015338	0.156863	8	43	id4532
ASXL1	20	31022599	DEL	A	-	Frame Shift Del	p.Q695fs	c.2084delA	NM_015338	0.19	9	38	id2275
ASXL1	20	31022643	DEL	G	-	Frame Shift Del	p.G710fs	c.2128delG	NM_015338	0.34	11	21	id14955
ASXL1	20	31022757	SNP	C	T	Nonsense Mutation	p.Q748*	c.2242C>T	NM_015338	0.3	9	21	id15381
ASXL1	20	31022784	SNP	C	T	Nonsense Mutation	p.Q757*	c.2269C>T	NM_015338	0.279412	19	49	id9809
ASXL1	20	31022793	SNP	C	T	Nonsense Mutation	p.Q760*	c.2278C>T	NM_015338	0.125	4	28	id12782
ASXL1	20	31022817	SNP	C	T	Nonsense Mutation	p.Q768*	c.2302C>T	NM_015338	0.1	5	45	id2116
ASXL1	20	31022817	SNP	C	T	Nonsense Mutation	p.Q768*	c.2302C>T	NM_015338	0.06	3	47	id11131
ASXL1	20	31022853	SNP	C	T	Nonsense Mutation	p.Q780*	c.2338C>T	NM_015338	0.292308	19	46	id9635
ASXL1	20	31022922	SNP	C	T	Nonsense Mutation	p.Q803*	c.2407C>T	NM_015338	0.112903	7	55	id10826
ASXL1	20	31022922	SNP	C	T	Nonsense Mutation	p.Q803*	c.2407C>T	NM_015338	0.296296	16	38	id11861
ASXL1	20	31022982	DEL	T	-	Frame Shift Del	p.L823fs	c.2467delT	NM_015338	0.08	10	111	id2961
ASXL1	20	31022983	SNP	T	G	Nonsense Mutation	p.L823*	c.2467delT	NM_015338	0.060403	9	140	id1808
ASXL1	20	31023271	INS	-	A	Frame Shift Ins	p.I919fs	c.2756_2757insA	NM_015338	0.22	18	64	id9984
ASXL1	20	31023354	SNP	G	T	Nonsense Mutation	p.E947*	c.2840A>G	NM_015338	0.08	6	69	id1754
ASXL1	20	31023381	DEL	CTTA	-	Frame Shift Del	p.L956fs	c.2866_2869delICT TA	NM_015338	0.31	25	56	id14964
ASXL1	20	31023528	DEL	T	-	Frame Shift Del	p.F1005fs	c.3013delT	NM_015338	0.1	11	98	id4445
ASXL1	20	31023554	DEL	C	-	Frame Shift Del	p.S1013fs	c.3039delC	NM_015338	0.12	11	81	id13638
ASXL1	20	31023625	SNP	G	A	Nonsense Mutation	p.W1037*	c.3110G>A	NM_015338	0.072581	9	115	id16502
ASXL1	20	31023702	SNP	C	T	Nonsense Mutation	p.Q1063*	c.3187C>T	NM_015338	0.052632	6	108	id3226
BCL11B	14	99641387	SNP	C	T	Misense Mutation	p.G596S	c.1786G>A	NM_138576	0.404255	19	28	id3170
BCOR	X	39922072	INS	-	G	Frame Shift Ins	p.H1367fs	c.4099_4100insC	NM_001123385	0.05	6	120	id14443
BCOR	X	39930272	DEL	CG	-	Frame Shift Del	p.S1064fs	c.3191_3192delCG	NM_001123385	0.1	12	109	id5900
BCOR	X	39933743	DEL	T	-	Frame Shift Del	p.S286fs	c.856delA	NM_001123385	0.43	12	16	id3439
BCOR	X	39934170	INS	-	T	Frame Shift Ins	p.N143fs	c.428_429insA	NM_001123385	0.03	6	185	id6419
BCORL1	X	129147623	SNP	C	G	Nonsense Mutation	p.S292*	c.875C>G	NM_021946	0.162791	7	36	id16590
BCORL1	X	129159270	SNP	C	T	Nonsense Mutation	p.R1332*	c.3994C>T	NM_021946	0.09375	6	58	id3227
BIRC3	11	102195867	INS	-	G	Frame Shift Ins	p.G209fs	c.627_628insG	NM_182962	0.04	7	163	id11974
BRAF	7	140453136	SNP	A	T	Misense Mutation	p.V600E	c.1799T>A	NM_004333	0.196429	11	45	id9117
BRAF	7	140481402	SNP	C	T	Misense Mutation	p.G469E	c.1406G>A	NM_004333	0.12973	24	161	id15992
BRCC3	X	154305490	SNP	C	T	Nonsense Mutation	p.R81*	c.241C>T	NM_024332	0.143791	22	131	id6955
BRCC3	X	154305490	SNP	C	T	Nonsense Mutation	p.R81*	c.241C>T	NM_024332	0.084337	7	76	id10444
BRCC3	X	154305514	SNP	C	T	Nonsense Mutation	p.R89*	c.265C>T	NM_024332	0.067416	6	83	id10536
BRCC3	X	154305514	SNP	C	T	Nonsense Mutation	p.R89*	c.265C>T	NM_024332	0.116667	7	53	id14164
BRCC3	X	154306935	SNP	G	A	Nonsense Mutation	p.W120*	c.360G>A	NM_024332	0.04	7	168	id8978
BRCC3	X	154306978	SNP	G	T	Splice Site	c.e5+1	c.403_splice	NM_024332	0.104348	12	103	id8963
CARD11	7	2976745	SNP	G	A	Misense Mutation	p.R423W	c.1267C>T	NM_032415	0.073171	3	38	id5921
CBL	11	119148921	SNP	T	C	Misense Mutation	p.C381R	c.1141T>C	NM_005188	0.122807	7	50	id4445
CBL	11	119148922	SNP	G	A	Misense Mutation	p.C381Y	c.1142G>A	NM_005188	0.34375	11	21	id15991
CBL	11	119148922	SNP	G	C	Misense Mutation	p.C381S	c.1142G>C	NM_005188	0.071429	3	39	id5319
CBL	11	119148924	SNP	A	G	Misense Mutation	p.C382E	c.1144A>G	NM_005188	0.064103	5	73	id3465
CBL	11	119148929	SNP	A	G	Misense Mutation	p.I383M	c.1149A>G	NM_005188	0.06383	6	88	id12878

Gene name	Chrom	Start position	Variant Type	Reference Allele	Variant Allele	Variant Classification	Protein Change	cDNA change	Accession	Variant allele fraction	Variant allele count	Reference allele count	ID
CBL	11	119148991	SNP	G	A	Missense_Mutation	p.C404Y	c.1211G>A	NM_005188	0.113333	17	133	id8673
CBL	11	119148991	SNP	G	A	Missense_Mutation	p.C404Y	c.1211G>A	NM_005188	0.072917	7	89	id15632
CBL	11	119148991	SNP	G	A	Missense_Mutation	p.C404Y	c.1211G>A	NM_005188	0.101562	13	115	id10499
CBL	11	119149235	SNP	G	A	Missense_Mutation	p.G415S	c.1243G>A	NM_005188	0.046875	6	122	id16015
CBL	11	119149235	SNP	G	A	Missense_Mutation	p.G415S	c.1243G>A	NM_005188	0.056	7	118	id15505
CBL	11	119149244	SNP	T	A	Missense_Mutation	p.F418I	c.1252T>A	NM_005188	0.082278	13	145	id3807
CBL	11	119149251	SNP	G	A	Missense_Mutation	p.R420Q	c.1259G>A	NM_005188	0.052632	6	108	id10937
CD58	1	117078587	SNP	C	T	Missense_Mutation	p.G210S	c.628G>A	NM_001779	0.368	46	79	id15637
CD79B	17	62006798	SNP	T	G	Missense_Mutation	p.Y196S	c.587A>C	NM_000626	0.086957	8	84	id6059
CNOT3	19	54646887	SNP	G	A	Missense_Mutation	p.E20K	c.58G>A	NM_014516	0.055556	12	204	id1662
CREBBP	16	3786067	DEL	T	-	Frame_Shift_Del	p.E1566fs	c.4698delA	NM_004380	0.03	7	236	id1443
CREBBP	16	3817721	DEL	T	-	Frame_Shift_Del	p.I1084fs	c.3250delA	NM_004380	0.05	8	149	id8258
CREBBP	16	3828056	INS	-	C	Frame_Shift_Ins	p.A690fs	c.2068_2069insG	NM_004380	0.03	6	182	id10673
CREBBP	16	3828080	INS	-	G	Frame_Shift_Ins	p.Q682fs	c.2044_2045insC	NM_004380	0.07	12	166	id3674
CREBBP	16	3843447	SNP	G	A	Nonsense_Mutation	p.R386*	c.1156C>T	NM_004380	0.054054	6	105	id14548
CREBBP	16	3860722	INS	-	G	Frame_Shift_Ins	p.Q286fs	c.856_857insC	NM_004380	0.08	10	122	id6570
CUX1	7	101559473	SNP	C	T	Nonsense_Mutation	p.Q37*	c.1124A>T	NM_181552	0.125	17	119	id3807
CUX1	7	101747648	SNP	C	T	Nonsense_Mutation	p.R147*	c.4430G>A	NM_181552	0.113636	15	117	id12656
CUX1	7	101837121	SNP	G	A	Splice_Site	c.e13-1	c.1077_splice	NM_181552	0.121212	4	29	id2276
DDX3X	X	41196685	INS	-	C	Frame_Shift_Ins	p.S24fs	c.70_71insC	NM_001356	0.09	18	190	id17182
DDX3X	X	41205659	DEL	CAGCAGTAT GTAT	-	Splice_Site	c.e13+1	c.1497_splice	NM_001356	0.17	20	97	id6943
DNMT3A	2	25457160	DEL	A	-	Frame_Shift_Del	p.F909fs	c.2727delT	NM_022552	0.04	7	183	id8553
DNMT3A	2	25457160	DEL	A	-	Frame_Shift_Del	p.F909fs	c.2727delT	NM_022552	0.05	7	142	id15476
DNMT3A	2	25457176	SNP	G	A	Missense_Mutation	p.P904L	c.2711C>T	NM_022552	0.164706	14	71	id3391
DNMT3A	2	25457176	SNP	G	A	Missense_Mutation	p.P904L	c.2711C>T	NM_022552	0.057143	4	66	id12744
DNMT3A	2	25457176	SNP	G	A	Missense_Mutation	p.P904L	c.2711C>T	NM_022552	0.054054	4	70	id8866
DNMT3A	2	25457176	SNP	G	A	Missense_Mutation	p.P904L	c.2711C>T	NM_022552	0.093333	7	68	id14640
DNMT3A	2	25457176	SNP	G	A	Missense_Mutation	p.P904L	c.2711C>T	NM_022552	0.068493	5	68	id15401
DNMT3A	2	25457176	SNP	G	A	Missense_Mutation	p.P904L	c.2711C>T	NM_022552	0.197368	15	61	id13118
DNMT3A	2	25457179	DEL	GC	-	Frame_Shift_Del	p.A903fs	c.2707_2708delGC	NM_022552	0.06	6	88	id15118
DNMT3A	2	25457181	INS	-	A	Frame_Shift_Ins	p.F902fs	c.2705_2706insT	NM_022552	0.06	7	118	id6054
DNMT3A	2	25457185	SNP	A	T	Missense_Mutation	p.L901H	c.2702T>A	NM_022552	0.083333	6	66	id14602
DNMT3A	2	25457207	INS	-	C	Frame_Shift_Ins	p.W893fs	c.2679_2680insG	NM_022552	0.08	6	67	id12337
DNMT3A	2	25457218	SNP	C	T	Missense_Mutation	p.G890D	c.2669G>A	NM_022552	0.0625	3	45	id1006
DNMT3A	2	25457238	INS	-	A	Frame_Shift_Ins	p.L883fs	c.2648_2649insT	NM_022552	0.09	7	68	id3225
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.078947	3	35	id16125
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.320755	17	36	id3816
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.290323	27	66	id3392
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.231707	19	63	id14954
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.381818	21	34	id6065
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.302632	23	53	id8970
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.068966	4	54	id12364
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.210526	8	30	id14900
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.06383	3	44	id15861
DNMT3A	2	25457242	SNP	C	G	Missense_Mutation	p.R882P	c.2645G>C	NM_022552	0.136986	10	63	id3327
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.081081	6	68	id6943
DNMT3A	2	25457242	SNP	C	G	Missense_Mutation	p.R882P	c.2645G>C	NM_022552	0.087719	5	52	id9681
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.069767	3	40	id10973
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.153846	8	44	id13747
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.089888	8	81	id6895
DNMT3A	2	25457242	SNP	C	G	Missense_Mutation	p.R882P	c.2645G>C	NM_022552	0.183099	13	58	id3247
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.142857	12	72	id4970
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.202899	14	55	id15710
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.492958	35	36	id6805
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.084507	6	65	id8646
DNMT3A	2	25457242	SNP	C	G	Missense_Mutation	p.R882P	c.2645G>C	NM_022552	0.190476	12	51	id12236
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.136364	12	76	id6767
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.072727	4	51	id10899
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.136364	9	57	id14751
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.258621	15	43	id15615
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.169492	10	49	id15630
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.061728	5	76	id1921
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.065217	3	43	id15600
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.377358	20	33	id17014
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.078431	4	47	id9232
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.072727	4	51	id13491
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.090909	7	70	id4533
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.089744	7	71	id5608
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.285714	12	30	id1519
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.155556	7	38	id15436
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.211538	11	41	id5422
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.057143	4	66	id2809
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.242105	23	72	id2743
DNMT3A	2	25457242	SNP	C	G	Missense_Mutation	p.R882P	c.2645G>C	NM_022552	0.107143	6	50	id5188
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.057143	4	66	id2698
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.113924	9	70	id12547
DNMT3A	2	25457242	SNP	C	T	Missense_Mutation	p.R882H	c.2645G>A	NM_022552	0.1	7	63	id2672
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.122449	6	43	id16177
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.076923	3	36	id16090
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.065574	4	57	id9965
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.075	3	37	id2320
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.102941	7	61	id15309
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.076923	3	36	id15857
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.135593	8	51	id13686
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.104167	5	43	id2059
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.226667	17	58	id3197
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.302326	13	30	id3170
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.065217	6	86	id4835
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.097561	8	74	id3681
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.205882	14	54	id14695
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.076923	5	60	id6592
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.117647	4	30	id1759
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C						

Gene name	Chrom	Start position	Variant Type	Reference Allele	Variant Allele	Variant Classification	Protein Change	cDNA change	Accession	Variant allele fraction	Variant allele count	Reference allele count	ID	
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.1	7	63	id5194	
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.372881	22	37	id14241	
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.175439	10	47	id10367	
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.052632	4	72	id714	
DNMT3A	2	25457243	SNP	G	A	Missense_Mutation	p.R882C	c.2644C>T	NM_022552	0.196078	10	41	id362	
DNMT3A	2	25457245	SNP	C	A	Missense_Mutation	p.S881I	c.2642G>T	NM_022552	0.09434	5	48	id5767	
DNMT3A	2	25457249	SNP	T	C	Missense_Mutation	p.M880V	c.2638A>G	NM_022552	0.266667	16	44	id2054	
DNMT3A	2	25457249	SNP	T	C	Missense_Mutation	p.M880V	c.2638A>G	NM_022552	0.061538	4	61	id8376	
DNMT3A	2	25457281	SNP	C	A	Missense_Mutation	p.G869V	c.2606G>T	NM_022552	0.083333	4	44	id3283	
DNMT3A	2	25457283	DEL	A	-	Frame_Shift_Del	p.F868fs	c.2604delIT	NM_022552	0.24	13	41	id9594	
DNMT3A	2	25458593	SNP	C	T	Nonsense_Mutation	p.W860*	c.2578T>C	NM_022552	0.169811	9	44	id15965	
DNMT3A	2	25458595	INS	-	T	Frame_Shift_Ins	p.L859fs	c.2575_2576delITT	NM_022552	0.07	8	111	id3141	
DNMT3A	2	25458595	SNP	A	G	Missense_Mutation	p.W860R	c.2578T>C	NM_022552	0.288889	13	32	id16986	
DNMT3A	2	25458595	SNP	A	G	Missense_Mutation	p.W860R	c.2578T>C	NM_022552	0.036232	5	133	id6619	
DNMT3A	2	25458597	DEL	AA	-	Frame_Shift_Del	p.L859fs	c.2575_2576delITT	NM_022552	0.07	6	85	id2249	
DNMT3A	2	25458602	DEL	G	-	Frame_Shift_Del	p.D857fs	c.2571delIC	NM_022552	0.1	7	62	id9931	
DNMT3A	2	25458622	DEL	AG	-	Frame_Shift_Del	p.V850fs	c.2550_2551delICT	NM_022552	0.14	11	66	id5711	
DNMT3A	2	25458635	DEL	CTGGT	-	Frame_Shift_Del	p.D845fs	CAG	c.2534_2538delAC	NM_022552	0.21	15	57	id17135
DNMT3A	2	25458653	INS	-	A	Frame_Shift_Ins	p.I840fs	c.2519_2520insT	NM_022552	0.14	12	71	id13494	
DNMT3A	2	25458692	DEL	G	-	Frame_Shift_Del	p.F827fs	c.2481delIC	NM_022552	0.16	18	93	id15890	
DNMT3A	2	25458695	SNP	C	T	Splice_Site	c.e22-1	c.2479_splice	NM_022552	0.071429	4	52	id15335	
DNMT3A	2	25458696	SNP	T	C	Splice_Site	c.e22-1	c.2479_splice	NM_022552	0.146667	11	64	id8979	
DNMT3A	2	25458696	SNP	T	C	Splice_Site	c.e22-1	c.2479_splice	NM_022552	0.051948	4	73	id15858	
DNMT3A	2	25458696	SNP	T	C	Splice_Site	c.e22-1	c.2479_splice	NM_022552	0.0625	4	60	id10900	
DNMT3A	2	25458696	SNP	T	C	Splice_Site	c.e22-1	c.2479_splice	NM_022552	0.103448	9	78	id11120	
DNMT3A	2	25459804	SNP	C	T	Splice_Site	c.e21+1	c.2478_splice	NM_022552	0.114286	4	31	id6005	
DNMT3A	2	25459804	SNP	C	T	Splice_Site	c.e21+1	c.2478_splice	NM_022552	0.166667	3	15	id17046	
DNMT3A	2	25459872	SNP	G	A	Missense_Mutation	p.P804L	c.2411C>T	NM_022552	0.18	9	41	id2965	
DNMT3A	2	25462011	SNP	G	T	Missense_Mutation	p.P799H	c.2396C>A	NM_022552	0.166667	4	20	id13451	
DNMT3A	2	25462015	DEL	G	-	Frame_Shift_Del	p.L798fs	c.2392delIC	NM_022552	0.19	10	44	id15563	
DNMT3A	2	25462020	SNP	C	A	Missense_Mutation	p.G796V	c.2387G>T	NM_022552	0.090909	4	40	id8952	
DNMT3A	2	25462022	SNP	C	T	Nonsense_Mutation	p.W795*	c.2385G>A	NM_022552	0.12069	7	51	id3421	
DNMT3A	2	25462053	DEL	A	-	Frame_Shift_Del	p.V785fs	c.2354delIT	NM_022552	0.21	13	50	id3033	
DNMT3A	2	25463171	SNP	C	G	Missense_Mutation	p.E774D	c.2322G>C	NM_022552	0.061699	8	123	id3224	
DNMT3A	2	25463171	SNP	C	T	Splice_Site	c.e19+1	c.2322_splice	NM_022552	0.487179	38	40	id12147	
DNMT3A	2	25463175	SNP	A	C	Missense_Mutation	p.L773R	c.2318T>G	NM_022552	0.03937	5	122	id12726	
DNMT3A	2	25463179	SNP	A	C	Missense_Mutation	p.F772V	c.2314T>G	NM_022552	0.038462	6	150	id17117	
DNMT3A	2	25463181	SNP	C	T	Missense_Mutation	p.R771Q	c.2312G>A	NM_022552	0.076923	7	84	id3722	
DNMT3A	2	25463181	SNP	C	T	Missense_Mutation	p.R771Q	c.2312G>A	NM_022552	0.05042	6	113	id1923	
DNMT3A	2	25463181	SNP	C	T	Missense_Mutation	p.R771Q	c.2312G>A	NM_022552	0.055556	8	136	id3034	
DNMT3A	2	25463181	SNP	C	T	Missense_Mutation	p.R771Q	c.2312G>A	NM_022552	0.044444	4	86	id1070	
DNMT3A	2	25463182	SNP	G	A	Nonsense_Mutation	p.R771*	c.2312G>A	NM_022552	0.051095	7	130	id16177	
DNMT3A	2	25463182	SNP	G	A	Nonsense_Mutation	p.R771*	c.2312G>A	NM_022552	0.223022	31	108	id9845	
DNMT3A	2	25463182	SNP	G	A	Nonsense_Mutation	p.R771*	c.2312G>A	NM_022552	0.191589	41	173	id6978	
DNMT3A	2	25463182	SNP	G	A	Nonsense_Mutation	p.R771*	c.2312G>A	NM_022552	0.088235	12	124	id10935	
DNMT3A	2	25463182	SNP	G	A	Nonsense_Mutation	p.R771*	c.2312G>A	NM_022552	0.065728	14	199	id4836	
DNMT3A	2	25463182	SNP	G	A	Nonsense_Mutation	p.R771*	c.2312G>A	NM_022552	0.04375	7	153	id2776	
DNMT3A	2	25463182	SNP	G	A	Nonsense_Mutation	p.R771*	c.2312G>A	NM_022552	0.112069	13	103	id14242	
DNMT3A	2	25463182	SNP	G	A	Nonsense_Mutation	p.R771*	c.2312G>A	NM_022552	0.224599	42	145	id88817	
DNMT3A	2	25463184	SNP	G	C	Missense_Mutation	p.S770W	c.2309C>G	NM_022552	0.09434	10	96	id15610	
DNMT3A	2	25463184	SNP	G	C	Missense_Mutation	p.S770W	c.2309C>G	NM_022552	0.046053	7	145	id15584	
DNMT3A	2	25463185	SNP	A	G	Missense_Mutation	p.S770P	c.2308T>C	NM_022552	0.03876	5	124	id14786	
DNMT3A	2	25463197	SNP	T	A	Nonsense_Mutation	p.K766*	c.2296A>T	NM_022552	0.049505	5	96	id8982	
DNMT3A	2	25463208	DEL	C	-	Frame_Shift_Del	p.G762fs	c.2285delIG	NM_022552	0.04	9	217	id2523	
DNMT3A	2	25463211	DEL	AT	-	Frame_Shift_Del	p.M761fs	c.2281_2282delIAT	NM_022552	0.28	38	96	id16600	
DNMT3A	2	25463212	SNP	T	C	Missense_Mutation	p.M761V	c.2281A>G	NM_022552	0.047847	10	199	id2963	
DNMT3A	2	25463213	DEL	G	-	Frame_Shift_Del	p.A760fs	c.2280delIC	NM_022552	0.07	14	198	id9680	
DNMT3A	2	25463228	SNP	A	C	Missense_Mutation	p.F755L	c.2265T>G	NM_022552	0.048276	7	138	id9897	
DNMT3A	2	25463229	SNP	A	G	Missense_Mutation	p.F755S	c.2264T>C	NM_022552	0.247619	26	79	id16185	
DNMT3A	2	25463229	SNP	A	G	Missense_Mutation	p.F755S	c.2264T>C	NM_022552	0.064286	9	131	id14878	
DNMT3A	2	25463229	SNP	A	G	Missense_Mutation	p.F755S	c.2264T>C	NM_022552	0.043103	5	111	id8657	
DNMT3A	2	25463230	SNP	A	T	Missense_Mutation	p.F755I	c.2263T>A	NM_022552	0.121622	9	65	id16013	
DNMT3A	2	25463230	SNP	A	T	Missense_Mutation	p.F755I	c.2263T>A	NM_022552	0.043478	6	132	id5386	
DNMT3A	2	25463231	DEL	G	-	Frame_Shift_Del	p.L754fs	c.2262delIC	NM_022552	0.34	41	78	id12579	
DNMT3A	2	25463232	SNP	A	T	Missense_Mutation	p.L754H	c.2261T>A	NM_022552	0.070796	8	105	id9280	
DNMT3A	2	25463235	SNP	C	T	Nonsense_Mutation	p.W753*	c.2258G>A	NM_022552	0.282258	35	89	id15952	
DNMT3A	2	25463236	DEL	AGA	-	In_Frame_Del	p.F752del	T	c.2255_2257delITC	NM_022552	0.04	6	142	id2365
DNMT3A	2	25463237	SNP	G	C	Missense_Mutation	p.F752L	c.2256C>G	NM_022552	0.0625	6	90	id8525	
DNMT3A	2	25463238	DEL	A	-	Frame_Shift_Del	p.F752fs	c.2255delIT	NM_022552	0.09	12	128	id3758	
DNMT3A	2	25463239	SNP	A	C	Missense_Mutation	p.F752V	c.2254T>G	NM_022552	0.093168	15	146	id12121	
DNMT3A	2	25463239	SNP	A	T	Missense_Mutation	p.F752I	c.2254T>A	NM_022552	0.165414	22	111	id13396	
DNMT3A	2	25463239	SNP	A	C	Missense_Mutation	p.F752V	c.2254T>G	NM_022552	0.047945	7	139	id8192	
DNMT3A	2	25463240	DEL	G	-	Frame_Shift_Del	p.F751fs	c.2253delIC	NM_022552	0.07	9	117	id3805	
DNMT3A	2	25463240	DEL	G	-	Frame_Shift_Del	p.F751fs	c.2253delIC	NM_022552	0.04	7	184	id6954	
DNMT3A	2	25463240	SNP	G	C	Missense_Mutation	p.F751L	c.2253C>G	NM_022552	0.3125	35	77	id6937	
DNMT3A	2	25463243	DEL	G	-	Frame_Shift_Del	p.P750fs	c.2250delIC	NM_022552	0.04	9	208	id4361	
DNMT3A	2	25463246	DNP	GC	AA	Missense_Mutation	p.R749L	c.2246_2247GC>TT	NM_022552	0.302326	26	60	id8792	
DNMT3A	2	25463247	SNP	C	T	Missense_Mutation	p.R749H	c.2246G>A	NM_022552	0.042857	6	134	id1756	
DNMT3A	2	25463267	INS	-	CG	Frame_Shift_Ins	p.R742fs	c.2225_2226insCG	NM_022552	0.23	21	72	id16702	
DNMT3A	2	25463283	SNP	A	T	Missense_Mutation	p.L737H	c.2210T>A	NM_022552	0.11236	10	79	id9422	
DNMT3A	2	25463283	SNP	A	G	Missense_Mutation	p.L737P	c.2210T>C	NM_022552	0.070175	12	159	id6671	
DNMT3A	2	25463283	SNP	A	C	Missense_Mutation	p.L737R	c.2210T>G	NM_022552	0.064103	5	73	id6440	
DNMT3A	2	25463286	SNP	C	T	Missense_Mutation	p.R736H	c.2207G>A	NM_022552	0.035971	5	134	id2222	
DNMT3A	2	25463286	SNP	C	T	Missense_Mutation	p.R736H	c.2207G>A	NM_022552	0.206897	12	46	id15683	
DNMT3A	2	25463286	SNP	C	T	Missense_Mutation	p.R736H	c.2207G>A	NM_022552	0.130841	14	93	id9230	
DNMT3A	2	25463286	SNP	C	G	Missense_Mutation	p.R736P	c.2207G>C	NM_022552	0.051282	6	111	id2964	
DNMT3A	2	25463286	SNP	C	T	Missense_Mutation	p.R736H	c.2207G>A	NM_022552	0.202703	15	59	id1522	
DNMT3A	2	25463287	SNP	G	A	Missense_Mutation	p.R736C	c.2206C>T	NM_022552	0.069767	6	80	id10935	
DNMT3A	2	25463287	DEL	G	-	Frame_Shift_Del	p.R736fs	c.2206delIC	NM_022552	0.04	7	151	id4834	
DNMT3A	2	25463287	SNP	G	A	Missense_Mutation	p.R736C	c.2206C>T	NM_022552	0.045455	5	105	id1400	
DNMT3A	2	25463287	SNP	G	A	Missense_Mutation	p.R736C	c.2206C>T	NM_022552	0.107143	9	75	id11823	
DNMT3A	2	25463287	SNP	G	A	Missense_Mutation	p.R736C	c.2206C>T	NM_022552	0.066667	4	56	id8297	
DNMT3A	2	25463288	INS	-	A	Frame_Shift_Ins	p.Y735fs	c.2204_2205insT	NM					

Gene name	Chrom	Start position	Variant Type	Reference Allele	Variant Allele	Variant Classification	Protein Change	cDNA change	Accession	Variant allele fraction	Variant allele count	Reference allele count	ID
DNMT3A	2	25463289	SNP	T	C	Missense_Mutation	p.Y735C	c.2204A>G	NM_022552	0.0625	6	90	id14999
DNMT3A	2	25463289	SNP	T	C	Missense_Mutation	p.Y735C	c.2204A>G	NM_022552	0.125	11	77	id12769
DNMT3A	2	25463289	SNP	T	C	Missense_Mutation	p.Y735C	c.2204A>G	NM_022552	0.15493	11	60	id8729
DNMT3A	2	25463289	SNP	T	C	Missense_Mutation	p.Y735C	c.2204A>G	NM_022552	0.104478	7	60	id6944
DNMT3A	2	25463289	SNP	T	C	Missense_Mutation	p.Y735C	c.2204A>G	NM_022552	0.064935	5	72	id2196
DNMT3A	2	25463289	SNP	T	C	Missense_Mutation	p.Y735C	c.2204A>G	NM_022552	0.068966	6	81	id13765
DNMT3A	2	25463289	SNP	T	G	Missense_Mutation	p.Y735S	c.2204A>C	NM_022552	0.044444	4	86	id3658
DNMT3A	2	25463289	SNP	T	C	Missense_Mutation	p.Y735C	c.2204A>G	NM_022552	0.14	7	43	id15574
DNMT3A	2	25463291	SNP	G	C	Missense_Mutation	p.F734L	c.2202C>G	NM_022552	0.045455	5	105	id2350
DNMT3A	2	25463292	SNP	A	C	Missense_Mutation	p.F734C	c.2201T>G	NM_022552	0.1	8	72	id11518
DNMT3A	2	25463296	INS	-	A	Frame_Shift_Ins	p.F732fs	c.2196_2197insT	NM_022552	0.09	9	88	id15312
DNMT3A	2	25463298	SNP	A	G	Missense_Mutation	p.F732S	c.2195T>C	NM_022552	0.084746	5	54	id8808
DNMT3A	2	25463298	DEL	AAG	-	In_Frame_Del	p.731_732FF>F T	c.2193_2195delCT	NM_022552	0.05	7	124	id3448
DNMT3A	2	25463298	SNP	A	G	Missense_Mutation	p.F732S	c.2195T>C	NM_022552	0.092105	7	69	id9932
DNMT3A	2	25463298	DEL	AAG	-	In_Frame_Del	p.731_732FF>F T	c.2193_2195delCT	NM_022552	0.04	6	155	id6823
DNMT3A	2	25463298	DEL	AAG	-	In_Frame_Del	p.731_732FF>F T	c.2190_2194delCT	NM_022552	0.21	19	73	id5924
DNMT3A	2	25463299	DEL	AGAAG	-	Frame_Shift_Del	p.L730fs	TCT	NM_022552	0.08	8	88	id8675
DNMT3A	2	25463300	SNP	G	T	Missense_Mutation	p.F731L	c.2193C>A	NM_022552	0.067797	4	55	id9875
DNMT3A	2	25463307	SNP	C	T	Missense_Mutation	p.R729Q	c.2186G>A	NM_022552	0.053763	5	88	id5320
DNMT3A	2	25463308	SNP	G	A	Missense_Mutation	p.R729W	c.2185C>T	NM_022552	0.098592	7	64	id2396
DNMT3A	2	25463308	SNP	G	A	Missense_Mutation	p.R729W	c.2185C>T	NM_022552	0.088889	4	41	id16068
DNMT3A	2	25463308	SNP	G	A	Missense_Mutation	p.R729W	c.2185C>T	NM_022552	0.098361	6	55	id15009
DNMT3A	2	25463308	SNP	G	A	Missense_Mutation	p.R729W	c.2185C>T	NM_022552	0.05	4	76	id8983
DNMT3A	2	25463308	SNP	G	C	Missense_Mutation	p.R729G	c.2185C>G	NM_022552	0.088608	7	72	id6003
DNMT3A	2	25463308	SNP	G	A	Missense_Mutation	p.R729W	c.2185C>T	NM_022552	0.051948	4	73	id2025
DNMT3A	2	25463308	SNP	G	A	Missense_Mutation	p.R729W	c.2185C>T	NM_022552	0.074074	6	75	id15128
DNMT3A	2	25463320	SNP	C	T	Splice_Site	c.e19-1	c.2174_splice	NM_022552	0.040404	4	95	id2319
DNMT3A	2	25463320	SNP	C	T	Splice_Site	c.e19-1	c.2174_splice	NM_022552	0.052632	4	72	id14459
DNMT3A	2	25463308	SNP	C	T	Splice_Site	c.e19-1	c.2174_splice	NM_022552	0.072464	5	64	id1939
DNMT3A	2	25463308	SNP	C	T	Splice_Site	c.e19-1	c.2174_splice	NM_022552	0.04902	5	97	id5800
DNMT3A	2	25463308	SNP	C	T	Splice_Site	c.e19-1	c.2174_splice	NM_022552	0.057143	6	99	id7308
DNMT3A	2	25463308	SNP	C	T	Splice_Site	c.e19-1	c.2174_splice	NM_022552	0.054054	6	105	id1242
DNMT3A	2	25463308	SNP	C	T	Splice_Site	c.e19-1	c.2174_splice	NM_022552	0.037594	10	256	id6108
DNMT3A	2	25463509	SNP	C	T	Splice_Site	c.e19-1	c.2174_splice	NM_022552	0.054795	4	69	id1343
DNMT3A	2	25463523	SNP	C	T	Missense_Mutation	p.R720H	c.2159G>A	NM_022552	0.068323	11	150	id8630
DNMT3A	2	25463529	DEL	G	-	Frame_Shift_Del	p.P718fs	c.2153delC	NM_022552	0.06	8	136	id12206
DNMT3A	2	25463532	SNP	T	A	Missense_Mutation	p.N717I	c.2150A>T	NM_022552	0.079812	17	196	id6708
DNMT3A	2	25463541	SNP	G	C	Missense_Mutation	p.S714C	c.2141C>G	NM_022552	0.036765	5	131	id16144
DNMT3A	2	25463541	SNP	G	C	Missense_Mutation	p.S714C	c.2141C>G	NM_022552	0.042553	4	90	id15164
DNMT3A	2	25463541	SNP	G	C	Missense_Mutation	p.S714C	c.2141C>G	NM_022552	0.055556	5	85	id14273
DNMT3A	2	25463554	SNP	A	T	Missense_Mutation	p.C710S	c.2128T>A	NM_022552	0.125926	17	118	id3426
DNMT3A	2	25463554	SNP	A	T	Missense_Mutation	p.C710S	c.2128T>A	NM_022552	0.045161	7	148	id4883
DNMT3A	2	25463568	SNP	A	G	Missense_Mutation	p.I705T	c.2114T>C	NM_022552	0.044444	4	86	id13787
DNMT3A	2	25463568	SNP	A	G	Missense_Mutation	p.I705T	c.2114T>C	NM_022552	0.103774	11	95	id8691
DNMT3A	2	25463571	SNP	A	C	Missense_Mutation	p.V704G	c.2111T>G	NM_022552	0.042373	5	113	id12386
DNMT3A	2	25463578	SNP	C	A	Missense_Mutation	p.D702Y	c.2104G>T	NM_022552	0.062069	9	136	id16127
DNMT3A	2	25463578	SNP	C	T	Missense_Mutation	p.D702N	c.2104G>A	NM_022552	0.139241	11	68	id8614
DNMT3A	2	25463583	DEL	G	-	Frame_Shift_Del	p.P700fs	c.2099delC	NM_022552	0.07	6	80	id16162
DNMT3A	2	25463583	SNP	G	C	Missense_Mutation	p.P700R	c.2099C>G	NM_022552	0.059524	5	79	id8775
DNMT3A	2	25463584	SNP	G	A	Missense_Mutation	p.P700S	c.2098C>T	NM_022552	0.338028	24	47	id15026
DNMT3A	2	25463586	SNP	C	T	Missense_Mutation	p.G699D	c.2096G>A	NM_022552	0.17284	14	67	id1719
DNMT3A	2	25463587	SNP	C	T	Missense_Mutation	p.G699S	c.2095G>A	NM_022552	0.116667	7	53	id3790
DNMT3A	2	25463587	SNP	C	T	Missense_Mutation	p.G699S	c.2095G>A	NM_022552	0.207207	23	88	id3196
DNMT3A	2	25463587	SNP	C	T	Missense_Mutation	p.G699S	c.2095G>A	NM_022552	0.226667	17	58	id12570
DNMT3A	2	25463596	SNP	G	A	Nonsense_Mutation	p.Q696*	c.2086C>T	NM_022552	0.275	44	116	id9013
DNMT3A	2	25463601	SNP	T	C	Splice_Site	c.e18-1	c.2083_splice	NM_022552	0.197183	14	57	id2365
DNMT3A	2	25463601	SNP	T	C	Splice_Site	c.e18-1	c.2083_splice	NM_022552	0.129032	8	54	id5795
DNMT3A	2	25464428	INS	-	A	Splice_Site	c.e17+1	c.2082_splice	NM_022552	0.28	11	29	id2306
DNMT3A	2	25464429	SNP	A	C	Splice_Site	c.e17+1	c.2082_splice	NM_022552	0.063492	4	59	id15906
DNMT3A	2	25464430	SNP	C	T	Splice_Site	c.e17+1	c.2082_splice	NM_022552	0.104167	5	43	id16591
DNMT3A	2	25464456	SNP	T	A	Missense_Mutation	p.D686V	c.2057A>T	NM_022552	0.122807	7	50	id9977
DNMT3A	2	25464457	SNP	C	A	Missense_Mutation	p.D686Y	c.2056G>T	NM_022552	0.083333	4	44	id12975
DNMT3A	2	25464459	SNP	C	T	Missense_Mutation	p.G685E	c.2054G>A	NM_022552	0.114286	4	31	id3840
DNMT3A	2	25464459	SNP	C	G	Missense_Mutation	p.G685A	c.2054G>C	NM_022552	0.083333	4	44	id11374
DNMT3A	2	25464460	SNP	C	T	Missense_Mutation	p.G685R	c.2053G>A	NM_022552	0.080645	5	57	id14857
DNMT3A	2	25464486	SNP	C	T	Missense_Mutation	p.R676Q	c.2027G>A	NM_022552	0.102041	5	44	id16161
DNMT3A	2	25464505	INS	-	G	Frame_Shift_Ins	p.S669fs	c.2007_2008insC	NM_022552	0.27	13	36	id16841
DNMT3A	2	25464534	SNP	T	C	Missense_Mutation	p.Y660C	c.1979A>G	NM_022552	0.075	3	37	id4444
DNMT3A	2	25464537	SNP	C	T	Missense_Mutation	p.R659H	c.1976G>A	NM_022552	0.238095	5	16	id16060
DNMT3A	2	25464537	SNP	C	T	Missense_Mutation	p.R659H	c.1976G>A	NM_022552	0.066667	3	42	id17165
DNMT3A	2	25464544	SNP	C	T	Missense_Mutation	p.V657M	c.1969G>A	NM_022552	0.136364	3	19	id14784
DNMT3A	2	25464544	SNP	C	T	Missense_Mutation	p.V657M	c.1969G>A	NM_022552	0.111111	3	24	id3950
DNMT3A	2	25464544	SNP	C	T	Missense_Mutation	p.V657M	c.1969G>A	NM_022552	0.117647	6	45	id8852
DNMT3A	2	25464547	SNP	G	A	Nonsense_Mutation	p.Q656*	c.1966C>T	NM_022552	0.125	4	28	id11568
DNMT3A	2	25464554	SNP	C	A	Missense_Mutation	p.L653F	c.1959G>T	NM_022552	0.195122	8	33	id2375
DNMT3A	2	25467022	SNP	A	G	Splice_Site	c.e15+1	c.1851_splice	NM_022552	0.196078	10	41	id13422
DNMT3A	2	25467023	SNP	C	T	Splice_Site	c.e15+1	c.1851_splice	NM_022552	0.0625	4	60	id6567
DNMT3A	2	25467032	SNP	G	A	Nonsense_Mutation	p.Q615*	c.1843C>T	NM_022552	0.074627	5	62	id14957
DNMT3A	2	25467065	DEL	G	-	Frame_Shift_Del	p.R604fs	c.1810delC	NM_022552	0.12	6	43	id17044
DNMT3A	2	25467073	SNP	C	T	Nonsense_Mutation	p.W601*	c.1802G>A	NM_022552	0.117647	8	60	id8648
DNMT3A	2	25467083	SNP	G	A	Nonsense_Mutation	p.R598*	c.1792C>T	NM_022552	0.225806	14	48	id16194
DNMT3A	2	25467083	SNP	G	A	Nonsense_Mutation	p.R598*	c.1792C>T	NM_022552	0.083333	3	33	id14858
DNMT3A	2	25467083	SNP	G	A	Nonsense_Mutation	p.R598*	c.1792C>T	NM_022552	0.086207	5	53	id2009
DNMT3A	2	25467083	SNP	G	A	Nonsense_Mutation	p.R598*	c.1792C>T	NM_022552	0.246154	16	49	id14719
DNMT3A	2	25467083	SNP	G	A	Nonsense_Mutation	p.R598*	c.1792C>T	NM_022552	0.06383	3	44	id13566
DNMT3A	2	25467083	SNP	G	A	Nonsense_Mutation	p.R598*	c.1792C>T	NM_022552	0.105263	8	68	id2962
DNMT3A	2	25467083	SNP	G	A	Nonsense_Mutation	p.R598*	c.1792C>T	NM_022552	0.069767	3	40	id11375
DNMT3A	2	25467111	DEL	G	-	Frame_Shift_Del	p.H588fs	c.1764delC	NM_022552	0.24	15	47	id15339
DNMT3A	2	25467132	SNP	C	T	Nonsense_Mutation	p.W581*	c.1741T>G	NM_022552	0.282051	11	28	id6869
DNMT3A	2	25467134	SNP	A	C	Missense_Mutation	p.W581G	c.1741T>G	NM_022552	0.115385	3	23	id16011
DNMT3A	2	25467407	SNP	A	C	Splice_Site	c.e14+1	c.1667_splice	NM_022552	0.090909	4	40	id3119
DNMT3A	2	25467408	SNP	C	A	Splice_Site	c.e14+1	c.1667_splice	NM_022552	0.0625	4	60	id8406
DNMT3A	2	25467408	SNP	C	G	Splice_Site	c.e14+1	c.1667_splice	NM_022552	0.215686	11		

Gene name	Chrom	Start position	Variant Type	Reference Allele	Variant Allele	Variant Classification	Protein Change	cDNA change	Accession	Variant allele fraction	Variant allele count	Reference allele count	ID
DNMT3A	2	25467428	SNP	C	T	Missense_Mutation	p.G550R	c.1648G>A	NM_022552	0.065217	3	43	id15052
DNMT3A	2	25467436	SNP	A	T	Missense_Mutation	p.L547H	c.1640T>A	NM_022552	0.129412	11	74	id14965
DNMT3A	2	25467436	SNP	A	C	Missense_Mutation	p.L547R	c.1640T>G	NM_022552	0.083333	3	33	id8962
DNMT3A	2	25467436	SNP	A	T	Missense_Mutation	p.L547H	c.1640T>A	NM_022552	0.107692	7	58	id12146
DNMT3A	2	25467437	SNP	G	A	Missense_Mutation	p.L547F	c.1639C>T	NM_022552	0.25	19	57	id9916
DNMT3A	2	25467437	SNP	G	A	Missense_Mutation	p.L547F	c.1639C>T	NM_022552	0.173077	9	43	id8674
DNMT3A	2	25467447	INS	-	C	Frame_Shift_Ins	p.G543fs	c.1628_1629insG	NM_022552	0.05	6	105	id2607
DNMT3A	2	25467448	SNP	C	G	Missense_Mutation	p.G543A	c.1628G>C	NM_022552	0.052632	4	72	id12180
DNMT3A	2	25467449	SNP	C	A	Missense_Mutation	p.G543C	c.1627G>T	NM_022552	0.140845	10	61	id5941
DNMT3A	2	25467449	SNP	C	T	Missense_Mutation	p.G543S	c.1627G>A	NM_022552	0.08	6	69	id10674
DNMT3A	2	25467465	SNP	G	T	Nonsense_Mutation	p.C537*	c.1611C>A	NM_022552	0.111111	11	88	id1774
DNMT3A	2	25467468	SNP	G	T	Nonsense_Mutation	p.Y536*	c.1608C>A	NM_022552	0.126582	10	69	id14901
DNMT3A	2	25467471	DEL	G	-	Frame_Shift_Del	p.S535fs	c.1605delC	NM_022552	0.22	21	74	id3357
DNMT3A	2	25467476	SNP	G	A	Nonsense_Mutation	p.Q534*	c.1600C>T	NM_022552	0.047059	4	81	id1471
DNMT3A	2	25467481	DEL	C	-	Frame_Shift_Del	p.G532fs	c.1595delG	NM_022552	0.06	7	112	id13672
DNMT3A	2	25467492	SNP	G	C	Nonsense_Mutation	p.Y528*	c.1582_1583delTA	NM_022552	0.078947	3	35	id16331
DNMT3A	2	25467493	DEL	TA	-	Frame_Shift_Del	p.Y528fs	c.1582_1583delTA	NM_022552	0.11	8	68	id7028
DNMT3A	2	25467496	SNP	T	G	Missense_Mutation	p.Q527P	c.1580A>C	NM_022552	0.123288	9	64	id2858
DNMT3A	2	25467497	SNP	G	A	Nonsense_Mutation	p.Q527*	c.1580A>C	NM_022552	0.160714	9	47	id15539
DNMT3A	2	25467503	DEL	CA	-	Frame_Shift_Del	p.C524fs	c.1572_1573delTG	NM_022552	0.12	6	43	id2056
DNMT3A	2	25467522	SNP	C	G	Splice_Site	c.e14-1	c.1555 splice	NM_022552	0.194805	15	62	id3359
DNMT3A	2	25467523	SNP	T	C	Splice_Site	c.e14-1	c.1555 splice	NM_022552	0.050505	5	94	id4534
DNMT3A	2	25468121	SNP	C	T	Splice_Site	c.e13+1	c.1554 splice	NM_022552	0.093023	8	78	id14835
DNMT3A	2	25468121	SNP	C	T	Splice_Site	c.e13+1	c.1554 splice	NM_022552	0.095238	4	38	id1870
DNMT3A	2	25468128	INS	-	T	Frame_Shift_Ins	p.N516fs	c.1547_1548insA	NM_022552	0.11	10	79	id12758
DNMT3A	2	25468145	SNP	C	A	Nonsense_Mutation	p.G511*	c.1531G>T	NM_022552	0.047619	6	120	id5891
DNMT3A	2	25468153	INS	-	G	Frame_Shift_Ins	p.L508fs	c.1522_1523insC	NM_022552	0.1	12	106	id14752
DNMT3A	2	25468165	DEL	A	-	Frame_Shift_Del	p.L504fs	c.1511delT	NM_022552	0.1	8	73	id2118
DNMT3A	2	25468181	DEL	T	-	Frame_Shift_Del	p.S499fs	c.1495delA	NM_022552	0.07	11	137	id5858
DNMT3A	2	25468182	DEL	C	-	Frame_Shift_Del	p.G498fs	c.1494delG	NM_022552	0.08	10	121	id3360
DNMT3A	2	25468186	SNP	C	T	Missense_Mutation	p.C497Y	c.1490G>A	NM_022552	0.104348	12	103	id9043
DNMT3A	2	25468202	SNP	C	G	Splice_Site	c.e13-1	c.1475 splice	NM_022552	0.04	4	96	id3227
DNMT3A	2	25468203	DEL	T	-	Splice_Site	c.e13-1	c.1475 splice	NM_022552	0.06	6	99	id4969
DNMT3A	2	25468887	SNP	A	C	Splice_Site	c.e13-1	c.1475 splice	NM_022552	0.060606	6	93	id2258
DNMT3A	2	25468899	DEL	C	-	Frame_Shift_Del	p.R488fs	c.1464delG	NM_022552	0.07	7	92	id7288
DNMT3A	2	25468899	DEL	C	-	Frame_Shift_Del	p.R488fs	c.1464delG	NM_022552	0.06	6	102	id1412
DNMT3A	2	25468907	SNP	T	A	Nonsense_Mutation	p.K486*	c.1456A>T	NM_022552	0.06	6	94	id15373
DNMT3A	2	25468919	SNP	C	A	Nonsense_Mutation	p.E482*	c.1444G>T	NM_022552	0.184211	14	62	id12235
DNMT3A	2	25468934	SNP	C	T	Splice_Site	c.e12-1	c.1430 splice	NM_022552	0.102041	5	44	id15827
DNMT3A	2	25468935	SNP	T	G	Splice_Site	c.e12-1	c.1430 splice	NM_022552	0.055556	5	85	id16096
DNMT3A	2	25469028	SNP	C	T	Splice_Site	c.e12-1	c.1430 splice	NM_022552	0.055556	6	102	id3449
DNMT3A	2	25469028	SNP	C	T	Splice_Site	c.e12-1	c.1430 splice	NM_022552	0.046154	6	124	id9784
DNMT3A	2	25469028	SNP	C	T	Splice_Site	c.e12-1	c.1430 splice	NM_022552	0.038095	8	202	id4837
DNMT3A	2	25469041	SNP	C	A	Nonsense_Mutation	p.E473*	c.1417G>T	NM_022552	0.089655	13	132	id6670
DNMT3A	2	25469048	DEL	A	-	Frame_Shift_Del	p.I470fs	c.1410delT	NM_022552	0.11	15	118	id16080
DNMT3A	2	25469063	DEL	G	-	Frame_Shift_Del	p.P465fs	c.1395delC	NM_022552	0.05	7	138	id9719
DNMT3A	2	25469083	SNP	T	A	Nonsense_Mutation	p.K459*	c.1375A>T	NM_022552	0.041096	9	210	id19125
DNMT3A	2	25469096	DEL	GG	-	Frame_Shift_Del	p.A454fs	c.1361_1362delCC	NM_022552	0.24	56	175	id33344
DNMT3A	2	25469154	DEL	ACTT	-	Frame_Shift_Del	p.E434fs	c.1301_1304delAA GT	NM_022552	0.06	24	360	id6788
DNMT3A	2	25469156	DEL	T	-	Frame_Shift_Del	p.E434fs	c.1301_1304delAA GT	NM_022552	0.03	6	190	id8483
DNMT3A	2	25469165	DEL	GGATTCTT	-	Frame_Shift_Del	p.E427fs	c.1281_1293delAG AGAAAGATCCC	NM_022552	0.09	19	198	id16176
DNMT3A	2	25469488	SNP	C	T	Splice_Site	c.e10+1	c.1279 splice	NM_022552	0.060976	5	77	id4443
DNMT3A	2	25469527	SNP	A	C	Missense_Mutation	p.F414C	c.1241T>G	NM_022552	0.07563	9	110	id9420
DNMT3A	2	25469527	SNP	A	C	Missense_Mutation	p.F414C	c.1241T>G	NM_022552	0.090909	11	110	id10738
DNMT3A	2	25469541	SNP	C	T	Nonsense_Mutation	p.W409*	c.1227G>A	NM_022552	0.117647	12	90	id6084
DNMT3A	2	25469541	SNP	C	T	Nonsense_Mutation	p.W409*	c.1227G>A	NM_022552	0.043103	5	111	id12467
DNMT3A	2	25469541	SNP	C	T	Nonsense_Mutation	p.W409*	c.1227G>A	NM_022552	0.079365	10	116	id1860
DNMT3A	2	25469542	DEL	CATT	-	Frame_Shift_Del	p.E408fs	c.1223_1226delAA TG	NM_022552	0.11	10	81	id14929
DNMT3A	2	25469553	DEL	GG	-	Frame_Shift_Del	p.P405fs	c.1214_1215delCC	NM_022552	0.08	9	108	id15015
DNMT3A	2	25469564	SNP	G	A	Nonsense_Mutation	p.Q402*	c.1204C>T	NM_022552	0.045752	7	146	id8920
DNMT3A	2	25469564	SNP	G	A	Nonsense_Mutation	p.Q402*	c.1204C>T	NM_022552	0.131868	12	79	id3510
DNMT3A	2	25469566	INS	-	T	Frame_Shift_Ins	p.V401fs	c.1201_1202insA	NM_022552	0.08	14	171	id2808
DNMT3A	2	25469607	SNP	G	CA	Frame_Shift_Ins	p.C387fs	c.1160_1161insTG	NM_022552	0.04	10	266	id8934
DNMT3A	2	25469613	INS	-	G	Frame_Shift_Ins	p.P385fs	c.1154delC	NM_022552	0.06	20	322	id16036
DNMT3A	2	25469614	DEL	G	-	Frame_Shift_Del	p.P385fs	c.1154delC	NM_022552	0.08	21	239	id8994
DNMT3A	2	25469614	DEL	G	-	Frame_Shift_Del	p.P385fs	c.1154delC	NM_022552	0.05	11	218	id6559
DNMT3A	2	25469625	DEL	C	-	Frame_Shift_Del	p.P381fs	c.1143delG	NM_022552	0.04	11	253	id15874
DNMT3A	2	25469625	DEL	C	-	Frame_Shift_Del	p.P381fs	c.1143delG	NM_022552	0.12	36	274	id15866
DNMT3A	2	25469646	SNP	C	G	Splice_Site	c.e10-1	c.1123 splice	NM_022552	0.060241	5	78	id12659
DNMT3A	2	25469647	SNP	T	C	Splice_Site	c.e10-1	c.1123 splice	NM_022552	0.072464	10	128	id13854
DNMT3A	2	25469919	SNP	C	T	Splice_Site	c.e9+1	c.1122 splice	NM_022552	0.115385	3	23	id6691
DNMT3A	2	25469919	SNP	C	T	Splice_Site	c.e9+1	c.1122 splice	NM_022552	0.290323	9	22	id16501
DNMT3A	2	25469922	SNP	G	A	Nonsense_Mutation	p.Q374*	c.1121A>C	NM_022552	0.133333	6	39	id17176
DNMT3A	2	25469922	SNP	G	A	Nonsense_Mutation	p.Q374*	c.1121A>C	NM_022552	0.190476	4	17	id11908
DNMT3A	2	25469945	SNP	C	T	Missense_Mutation	p.R366H	c.1097G>A	NM_022552	0.083333	3	33	id10859
DNMT3A	2	25469945	SNP	C	T	Missense_Mutation	p.R366H	c.1097G>A	NM_022552	0.078947	3	35	id10082
DNMT3A	2	25469946	SNP	G	C	Missense_Mutation	p.R366G	c.1096C>G	NM_022552	0.073171	3	38	id4535
DNMT3A	2	25469976	SNP	G	A	Nonsense_Mutation	p.Q356*	c.1066C>T	NM_022552	0.092593	5	49	id9595
DNMT3A	2	25469976	SNP	G	A	Nonsense_Mutation	p.Q356*	c.1066C>T	NM_022552	0.087912	8	83	id8990
DNMT3A	2	25469984	DEL	G	-	Frame_Shift_Del	p.A353fs	c.1058delC	NM_022552	0.21	14	52	id15877
DNMT3A	2	25470029	INS	-	G	Splice_Site	c.e9-1	c.1015 splice	NM_022552	0.41	9	13	id12897
DNMT3A	2	25470467	DEL	A	-	Frame_Shift_Del	p.F336fs	c.1007delT	NM_022552	0.03	8	246	id12009
DNMT3A	2	25470480	SNP	C	A	Nonsense_Mutation	p.G332*	c.994G>T	NM_022552	0.078652	7	82	id6025
DNMT3A	2	25470484	SNP	C	T	Nonsense_Mutation	p.W330*	c.989_991delGGT	NM_022552	0.066038	7	99	id5896
DNMT3A	2	25470485	SNP	C	T	Nonsense_Mutation	p.W330*	c.989_991delGGT	NM_022552	0.064516	6	87	id12786
DNMT3A	2	25470493	SNP	C	T	Nonsense_Mutation	p.W327*	c.981G>A	NM_022552	0.077519	10	119	id5022
DNMT3A	2	25470493	SNP	C	T	Nonsense_Mutation	p.W327*	c.981G>A	NM_022552	0.049587	6	115	id928
DNMT3A	2	25470494	SNP	C	T	Nonsense_Mutation	p.W327*	c.981G>A	NM_022552	0.103448	12	104	id15000
DNMT3A	2	25470497	SNP	C	T	Missense_Mutation	p.R326H	c.977G>A	NM_022552	0.082569	9	100	id9718
DNMT3A	2	25470497	SNP	C	T	Missense_Mutation	p.R326H	c.977G>A	NM_022552	0.044944	4	85	id12336
DNMT3A	2	25470497	SNP	C	A	Missense_Mutation	p.R326L	c.977G>T	NM_022552	0.25	21	63	id4839
DNMT3A	2	25470497	SNP	C	T	Missense_Mutation	p.R326H	c.977G>A	NM_022552	0.059524	5	79	id12997
DNMT3A	2	25470497	SNP	C	T	Missense_Mutation	p.R326H	c.977G>A	NM_022552	0.162602	20	103	id7161
DNMT3A	2	25470498	SNP	G	A								

Gene name	Chrom	Start position	Variant Type	Reference Allele	Variant Allele	Variant Classification	Protein Change	cDNA change	Accession	Variant allele fraction	Variant allele count	Reference allele count	ID
DNMT3A	2	25470498	SNP	G	A	Missense_Mutation	p.R326C	c.976C>T	NM_022552	0.045455	4	84	id2197
DNMT3A	2	25470498	SNP	G	T	Missense_Mutation	p.R326S	c.976C>A	NM_022552	0.229885	20	67	id17118
DNMT3A	2	25470498	SNP	G	A	Missense_Mutation	p.R326C	c.976C>T	NM_022552	0.044944	4	85	id4694
DNMT3A	2	25470498	SNP	G	A	Missense_Mutation	p.R326C	c.976C>T	NM_022552	0.04902	5	97	id1859
DNMT3A	2	25470498	SNP	G	A	Missense_Mutation	p.R326C	c.976C>T	NM_022552	0.193548	18	75	id2999
DNMT3A	2	25470498	SNP	G	A	Missense_Mutation	p.R326C	c.976C>T	NM_022552	0.045455	4	84	id13251
DNMT3A	2	25470498	SNP	G	A	Missense_Mutation	p.R326C	c.976C>T	NM_022552	0.04878	4	78	id12179
DNMT3A	2	25470498	SNP	G	A	Missense_Mutation	p.R326C	c.976C>T	NM_022552	0.236842	27	87	id4044
DNMT3A	2	25470516	SNP	G	A	Nonsense_Mutation	p.R320*	c.958C>T	NM_022552	0.156863	8	43	id2402
DNMT3A	2	25470516	SNP	G	A	Nonsense_Mutation	p.R320*	c.958C>T	NM_022552	0.333333	21	42	id14930
DNMT3A	2	25470516	SNP	G	A	Nonsense_Mutation	p.R320*	c.958C>T	NM_022552	0.051282	6	111	id15764
DNMT3A	2	25470516	SNP	G	A	Nonsense_Mutation	p.R320*	c.958C>T	NM_022552	0.067416	6	83	id11418
DNMT3A	2	25470516	SNP	G	A	Nonsense_Mutation	p.R320*	c.958C>T	NM_022552	0.054054	6	105	id13315
DNMT3A	2	25470516	SNP	G	A	Nonsense_Mutation	p.R320*	c.958C>T	NM_022552	0.102804	11	96	id1472
DNMT3A	2	25470516	SNP	G	A	Nonsense_Mutation	p.R320*	c.958C>T	NM_022552	0.058824	5	80	id15130
DNMT3A	2	25470516	SNP	G	A	Nonsense_Mutation	p.R320*	c.958C>T	NM_022552	0.061404	7	107	id17185
DNMT3A	2	25470516	SNP	G	A	Nonsense_Mutation	p.R320*	c.958C>T	NM_022552	0.112676	8	63	id12510
DNMT3A	2	25470532	SNP	C	T	Nonsense_Mutation	p.W314*	c.942G>A	NM_022552	0.05	4	76	id14163
DNMT3A	2	25470555	SNP	G	A	Missense_Mutation	p.P307S	c.919C>T	NM_022552	0.074627	5	62	id5900
DNMT3A	2	25470556	SNP	C	T	Nonsense_Mutation	p.W306*	c.916T>C	NM_022552	0.060241	5	78	id10739
DNMT3A	2	25470559	SNP	C	T	Nonsense_Mutation	p.W305*	c.915G>A	NM_022552	0.057971	4	65	id5942
DNMT3A	2	25470559	SNP	C	T	Nonsense_Mutation	p.W305*	c.915G>A	NM_022552	0.045977	4	83	id11145
DNMT3A	2	25470560	SNP	C	T	Nonsense_Mutation	p.W305*	c.915G>A	NM_022552	0.326087	15	31	id3438
DNMT3A	2	25470583	SNP	C	T	Nonsense_Mutation	p.W297*	c.889T>G	NM_022552	0.056075	6	101	id12120
DNMT3A	2	25470594	DEL	C	-	Frame_Shift_Del	p.E294fs	c.880delG	NM_022552	0.2	23	90	id8730
DNMT3A	2	25470594	SNP	C	A	Nonsense_Mutation	p.E294*	c.880delG	NM_022552	0.186047	16	70	id1768
DNMT3A	2	25470904	SNP	A	C	Splice_Site	c.e7+1	c.855_splice	NM_022552	0.097561	4	37	id4692
DNMT3A	2	25470905	SNP	C	T	Splice_Site	c.e7+1	c.855_splice	NM_022552	0.222222	10	35	id11641
DNMT3A	2	25470905	SNP	C	T	Splice_Site	c.e7+1	c.855_splice	NM_022552	0.081967	5	56	id16034
DNMT3A	2	25470905	SNP	C	A	Splice_Site	c.e7+1	c.855_splice	NM_022552	0.121212	4	29	id6449
DNMT3A	2	25470955	INS	-	C	Frame_Shift_Ins	p.A269fs	c.805_806insG	NM_022552	0.13	8	52	id1868
DNMT3A	2	25470982	DEL	GT	-	Frame_Shift_Del	p.T260fs	c.778_779delAC	NM_022552	0.06	13	192	id3379
DNMT3A	2	25471011	DEL	G	-	Frame_Shift_Del	p.P250fs	c.750delC	NM_022552	0.25	27	80	id15727
DNMT3A	2	25471016	SNP	G	A	Nonsense_Mutation	p.Q249*	c.745C>T	NM_022552	0.147059	15	87	id12637
DNMT3A	2	25471016	SNP	G	A	Nonsense_Mutation	p.Q249*	c.745C>T	NM_022552	0.047619	6	120	id5252
DNMT3A	2	25471030	DEL	G	-	Frame_Shift_Del	p.P244fs	c.731delC	NM_022552	0.14	15	92	id5902
DNMT3A	2	25471030	DEL	G	-	Frame_Shift_Del	p.P244fs	c.731delC	NM_022552	0.06	9	147	id14666
DNMT3A	2	25471039	DEL	T	-	Frame_Shift_Del	p.E241fs	c.722delA	NM_022552	0.06	7	106	id3226
DNMT3A	2	25471040	SNP	C	A	Nonsense_Mutation	p.E241*	c.722delA	NM_022552	0.082474	8	89	id6851
DNMT3A	2	25471048	INS	-	TC	Frame_Shift_Ins	p.K238fs	c.712_713insGA	NM_022552	0.06	8	126	id9185
DNMT3A	2	25471052	INS	-	A	Frame_Shift_Ins	p.S236fs	c.708_709insT	NM_022552	0.05	6	121	id2345
DNMT3A	2	25471069	DEL	T	-	Frame_Shift_Del	p.Q231fs	c.692delA	NM_022552	0.07	8	112	id17166
DNMT3A	2	25471082	DEL	C	-	Frame_Shift_Del	p.V227fs	c.679delG	NM_022552	0.26	18	51	id11008
DNMT3A	2	25523009	SNP	G	A	Splice_Site	c.e3+1	c.177_splice	NM_022552	0.407407	11	16	id8961
EP300	22	41523642	INS	-	C	Frame_Shift_Ins	p.R353fs	c.1058_1059insC	NM_001429	0.08	21	231	id4590
ETV6	12	12022502	DEL	C	-	Frame_Shift_Del	p.S203fs	c.608delC	NM_001987	0.03	6	194	id3843
EZH2	7	148506461	SNP	C	T	Missense_Mutation	p.R679H	c.2036G>A	NM_001203247	0.047297	7	141	id15352
EZH2	7	148515097	INS	-	T	Frame_Shift_Ins	p.S366fs	c.1096_1097insA	NM_001203247	0.04	8	177	id6929
FAM46C	1	118165768	INS	-	C	Frame_Shift_Ins	p.L93fs	c.278_279insC	NM_017709	0.14	20	122	id16047
FBXW7	4	153268225	INS	-	A	Splice_Site	c.e4-1	c.585_splice	NM_033632	0.12	6	44	id6058
FLT3	13	28608281	SNP	A	G	Missense_Mutation	p.V592A	c.1775T>C	NM_004119	0.066667	7	98	id6574
FOXP1	3	71026170	INS	-	T	Frame_Shift_Ins	p.K484fs	c.1451_1452insA	NM_032682	0.05	7	134	id15631
GNAS	20	57484420	SNP	C	G	Missense_Mutation	p.R844G	c.2530C>G	NM_016592	0.260504	31	88	id12468
GNAS	20	57484420	SNP	C	T	Missense_Mutation	p.R844C	c.2530C>T	NM_016592	0.128205	10	68	id9846
GNAS	20	57484420	SNP	C	A	Missense_Mutation	p.R844S	c.2530C>A	NM_016592	0.080645	5	57	id16703
GNAS	20	57484421	SNP	G	A	Missense_Mutation	p.R844H	c.2531G>A	NM_016592	0.066667	5	70	id6830
GNAS	20	57484421	SNP	G	A	Missense_Mutation	p.R844H	c.2531G>A	NM_016592	0.121951	5	36	id12269
GNAS	20	57484421	SNP	G	A	Missense_Mutation	p.R844H	c.2531G>A	NM_016592	0.054348	5	87	id4939
GNAS	20	57484421	SNP	G	A	Missense_Mutation	p.R844H	c.2531G>A	NM_016592	0.05	4	76	id5860
GNAS	20	57484421	SNP	G	A	Missense_Mutation	p.R844H	c.2531G>A	NM_016592	0.155556	7	38	id15544
GNB1	1	1747227	SNP	C	A	Missense_Mutation	p.K57N	c.171G>T	NM_002074	0.098039	5	46	id3168
GNB1	1	1747227	SNP	C	A	Missense_Mutation	p.K57N	c.171G>T	NM_002074	0.131579	5	33	id13397
GNB1	1	1747228	SNP	T	A	Missense_Mutation	p.K57M	c.170A>T	NM_002074	0.096154	5	47	id12470
GNB1	1	1747228	SNP	T	A	Missense_Mutation	p.K57M	c.170A>T	NM_002074	0.083333	5	55	id4360
GNB1	1	1747228	SNP	T	A	Missense_Mutation	p.K57M	c.170A>T	NM_002074	0.211538	11	41	id8456
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.214286	9	33	id11645
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.121212	4	29	id8807
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.285714	10	25	id11639
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.090909	3	30	id3437
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.16	4	21	id3822
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.088235	3	31	id8979
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.321429	9	19	id5020
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.065217	3	43	id4971
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.384615	15	24	id10936
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.4375	14	18	id8647
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.081633	4	45	id4772
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.068966	4	54	id15602
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.075	3	37	id9339
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.278689	17	44	id6556
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.208333	10	38	id12008
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.130435	6	40	id1659
GNB1	1	1747229	SNP	T	C	Missense_Mutation	p.K57E	c.169A>G	NM_002074	0.104167	5	43	id4207
HIST1H1C	6	26056332	SNP	T	A	Nonsense_Mutation	p.K109*	c.326A>G	NM_005319	0.055215	9	154	id4838
IDH2	15	90631934	SNP	C	T	Missense_Mutation	p.R140Q	c.419G>A	NM_002168	0.094118	8	77	id12629
IDH2	15	90631934	SNP	C	T	Missense_Mutation	p.R140Q	c.419G>A	NM_002168	0.043011	4	89	id11438
IDH2	15	90631934	SNP	C	T	Missense_Mutation	p.R140Q	c.419G>A	NM_002168	0.126437	11	76	id562
IKZF1	7	50468196	SNP	C	A	Nonsense_Mutation	p.Y477*	c.1431C>A	NM_006060	0.071429	3	39	id2011
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.05814	5	81	id16145
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.217742	27	97	id11641
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.27193	31	83	id9951
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.082569	9	100	id9952
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.058824	8	128	id3449
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.225806	21	72	id2348
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.287356	25	62	id15971
JAK2	9	5073770	SNP	G	T								

Gene name	Chrom	Start position	Variant Type	Reference Allele	Variant Allele	Variant Classification	Protein Change	cDNA change	Accession	Variant allele fraction	Variant allele count	Reference allele count	ID
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.041237	4	93	id12296
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.420561	45	62	id13636
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.040816	4	94	id9457
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.072289	12	154	id4774
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.888	111	14	id5826
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.089888	8	81	id16987
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.062992	8	119	id15181
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.252252	28	83	id11973
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.043956	4	87	id15486
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.071429	8	104	id16838
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.097015	13	121	id13371
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.048544	10	196	id1470
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.241935	30	94	id12573
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.073171	9	114	id6354
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.05618	5	84	id15417
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.048193	8	158	id1341
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.438462	57	73	id5286
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.090909	9	90	id3877
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.111111	9	72	id12878
JAK2	9	5073770	SNP	G	T	Missense_Mutation	p.V617F	c.1849G>T	NM_004972	0.051095	7	130	id428
JAK2	19	17949108	SNP	C	T	Missense_Mutation	p.M511I	c.1533G>A	NM_000215	0.225108	52	179	id1146
JARID2	6	15501273	INS	-	G	Frame_Shift_Ins	p.L694fs	c.2081_2082insG	NM_004973	0.06	6	90	id1308
KDM6A	X	44879925	SNP	C	T	Nonsense_Mutation	p.R172*	c.514C>T	NM_021140	0.647059	55	30	id12922
KDM6A	X	44942739	DEL	C	-	Frame_Shift_Del	p.P1107fs	c.3319delC	NM_021140	0.27	8	22	id3440
KIT	4	55594197	SNP	C	T	Missense_Mutation	p.R634W	c.1900C>T	NM_000222	0.113636	10	78	id3248
KLHL6	3	183273150	SNP	T	A	Missense_Mutation	p.R98W	c.292A>T	NM_130446	0.061728	5	76	id14548
KRAS	12	25378561	SNP	G	A	Missense_Mutation	p.A146V	c.437C>T	NM_033360	0.083333	4	44	id15632
KRAS	12	25378647	SNP	T	G	Missense_Mutation	p.K117N	c.351A>C	NM_033360	0.2625	21	59	id13666
KRAS	12	25398285	SNP	C	A	Missense_Mutation	p.G12C	c.34G>T	NM_033360	0.125	3	21	id4833
LCU7L2	7	139083380	DEL	C	-	Frame_Shift_Del	p.D64fs	c.192delC	NM_016019	0.1	18	160	id7124
LCU7L2	7	139086898	SNP	C	T	Nonsense_Mutation	p.Q91*	c.271C>T	NM_016019	0.119048	5	37	id5952
LCU7L2	7	139097301	SNP	C	T	Nonsense_Mutation	p.R262*	c.785G>A	NM_016019	0.128788	17	115	id4691
MLL2	19	36213944	INS	-	G	Frame_Shift_Ins	p.R924fs	c.2770_2771insG	NM_003482	0.05	6	126	id12338
MLL2	12	49427936	INS	-	T	Frame_Shift_Ins	p.K3551fs	c.10653_10654insA	NM_003482	0.13	16	111	id7288
MLL2	12	49434301	SNP	G	A	Nonsense_Mutation	p.Q2418*	c.7252C>T	NM_003482	0.069767	3	40	id17045
MLL2	12	49444933	DEL	G	-	Frame_Shift_Del	p.R845fs	c.2533delC	NM_003482	0.03	7	215	id8776
MPL	1	43814993	DEL	CTG	-	In_Frame_Del	p.L513del	c.1528_1530delCT	NM_005373	0.03	6	190	id15871
MPL	1	43815009	SNP	G	T	Missense_Mutation	p.W515L	c.1544G>T	NM_005373	0.632653	62	36	id3439
MYD88	3	38182641	SNP	T	C	Missense_Mutation	p.L273P	c.818T>C	NM_001172567	0.10989	10	81	id4362
MYD88	3	38182641	SNP	T	C	Missense_Mutation	p.L273P	c.818T>C	NM_001172567	0.128205	10	68	id12559
NOTCH1	9	139400000	SNP	C	A	Nonsense_Mutation	p.E1450*	c.4348G>A	NM_017617	0.06383	3	44	id12729
NOTCH1	9	139413957	DEL	T	-	Frame_Shift_Del	p.N268fs	c.803delA	NM_017617	0.02	6	272	id966
NOTCH2	1	120467977	SNP	C	A	Nonsense_Mutation	p.E1488*	c.4462G>T	NM_024408	0.044944	12	255	id11553
NOTCH2	1	120611991	SNP	C	T	Nonsense_Mutation	p.W10*	c.30G>A	NM_024408	0.090909	3	30	id12422
NRAS	1	115256528	SNP	T	G	Missense_Mutation	p.Q61H	c.183A>C	NM_002524	0.107914	15	124	id15579
NRAS	1	115256529	SNP	T	C	Missense_Mutation	p.Q61R	c.182A>G	NM_002524	0.045198	8	169	id2857
NRAS	1	115256530	SNP	G	T	Missense_Mutation	p.Q61K	c.181C>A	NM_002524	0.055276	11	188	id15859
NRAS	1	115256532	SNP	C	T	Missense_Mutation	p.G60E	c.179G>A	NM_002524	0.049296	7	135	id7257
NRAS	1	115258744	SNP	C	T	Missense_Mutation	p.G13D	c.38G>A	NM_002524	0.065421	14	200	id16047
NRAS	1	115258745	SNP	C	G	Missense_Mutation	p.G13R	c.37G>C	NM_002524	0.057971	12	195	id5915
PDS52	6	107655446	DEL	G	-	Frame_Shift_Del	p.N129fs	c.387delC	NM_020381	0.52	47	43	id4938
PHF6	X	133551280	SNP	G	T	Nonsense_Mutation	p.G306*	c.916G>T	NM_001015877	0.161765	11	57	id9810
PIK3CA	3	178916726	SNP	G	A	Missense_Mutation	p.R38H	c.113G>A	NM_006218	0.39759	33	50	id9419
PRDM1	6	106553280	INS	-	C	Frame_Shift_Ins	p.L415fs	c.1245_1246insC	NM_001198	0.5	39	39	id2010
PRDM1	6	106553536	INS	-	C	Frame_Shift_Ins	p.A501fs	c.1501_1502insC	NM_001198	0.09	6	58	id6622
PRDM1	6	106553596	INS	-	C	Frame_Shift_Ins	p.T521fs	c.1561_1562insC	NM_001198	0.15	6	35	id6724
PRPF40B	12	50027850	SNP	C	T	Nonsense_Mutation	p.Q241*	c.721C>T	NM_001031698	0.22	11	39	id11569
PTPN11	12	112888172	SNP	A	G	Missense_Mutation	p.Y63C	c.188A>G	NM_002834	0.040404	4	95	id15878
PTPN11	12	112888172	SNP	A	G	Missense_Mutation	p.Y63C	c.188A>G	NM_002834	0.382609	44	71	id4772
PTPN11	12	112888189	SNP	G	C	Missense_Mutation	p.E69Q	c.205G>C	NM_002834	0.20354	23	90	id9785
RAD21	8	117862942	INS	-	T	Frame_Shift_Ins	p.I512fs	c.1534_1535insA	NM_006265	0.08	7	76	id16143
RAD21	8	117863003	SNP	G	A	Nonsense_Mutation	p.Q492*	c.1474C>T	NM_006265	0.04878	4	78	id1013
RAD21	8	117864908	DEL	GGTCTTCTG	-	Frame_Shift_Del	p.P398fs	c.1192_1201delCC	NM_006265	0.07	6	79	id8977
RAD21	8	117869618	DEL	TAGGAGGTT	-	Frame_Shift_Del	p.S189fs	c.566_576delCTAA	NM_006265	0.17	13	64	id16332
RAD21	8	117875449	SNP	C	T	Missense_Mutation	p.R65Q	c.194G>A	NM_006265	0.234483	34	111	id3404
RAD21	8	117875476	SNP	G	C	Nonsense_Mutation	p.S56*	c.167C>G	NM_006265	0.254777	40	117	id12270
RIT1	1	155874194	INS	-	A	Frame_Shift_Ins	p.R112fs	c.336_337insT	NM_006912	0.07	7	96	id8573
RPS15	19	1440427	SNP	C	G	Missense_Mutation	p.A135G	c.404C>G	NM_001018	0.277778	25	65	id3771
SETD2	3	47084145	INS	-	G	Frame_Shift_Ins	p.P2381fs	c.7143_7144insC	NM_014159	0.04	7	177	id7288
SETD2	3	47103730	DEL	TTTAT	-	Frame_Shift_Del	p.N2138fs	c.6413_6417delAAT	NM_014159	0.05	18	356	id12181
SETD2	3	47125328	INS	-	G	Frame_Shift_Ins	p.Q2048fs	c.6142_6143insC	NM_014159	0.02	6	259	id1369
SETD2	3	47158225	SNP	G	A	Nonsense_Mutation	p.R1492*	c.4474C>T	NM_014159	0.142857	13	78	id7102
SETDB1	1	150933381	INS	-	C	Frame_Shift_Ins	p.H948fs	c.2843_2844insC	NM_001145415	0.03	6	193	id7288
SF1	11	64533489	INS	-	G	Frame_Shift_Ins	p.Q574fs	c.1720_1721insC	NM_004630	0.11	6	47	id6458
SF1	11	64536811	SNP	C	G	Splice_Site	c.e7-1	c.664_splice	NM_004630	0.104167	5	43	id15592
SF3A1	22	30737857	INS	-	G	Frame_Shift_Ins	p.P298fs	c.894_895insC	NM_005877	0.05	7	143	id7140
SF3B1	2	198266489	SNP	C	T	Missense_Mutation	p.E783K	c.2347G>A	NM_012433	0.075269	7	86	id4885
SF3B1	2	198266611	SNP	C	T	Missense_Mutation	p.G742D	c.2225G>A	NM_012433	0.070175	4	53	id8807
SF3B1	2	198266834	SNP	T	C	Missense_Mutation	p.K700E	c.2098A>G	NM_012433	0.213333	16	59	id16099
SF3B1	2	198266834	SNP	T	C	Missense_Mutation	p.K700E	c.2098A>G	NM_012433	0.1	7	63	id16100
SF3B1	2	198266834	SNP	T	C	Missense_Mutation	p.K700E	c.2098A>G	NM_012433	0.230769	12	40	id12421
SF3B1	2	198266834	SNP	T	C	Missense_Mutation	p.K700E	c.2098A>G	NM_012433	0.058824	4	64	id3358
SF3B1	2	198266834	SNP	T	C	Missense_Mutation	p.K700E	c.2098A>G	NM_012433	0.140187	15	92	id3345
SF3B1	2	198266834	SNP	T	C	Missense_Mutation	p.K700E	c.2098A>G	NM_012433	0.070175	4	53	id14902
SF3B1	2	198266834	SNP	T	C	Missense_Mutation	p.K700E	c.2098A>G	NM_012433	0.133333	10	65	id3195
SF3B1	2	198266834	SNP	T	C	Missense_Mutation	p.K700E	c.2098A>G	NM_012433	0.075	6	74	id9423
SF3B1	2	198266834	SNP	T	C	Missense_Mutation	p.K700E	c.2098A>G	NM_012433	0.113208	6	47	id15231
SF3B1	2	198266834	SNP	T	C	Missense_Mutation	p.K700E	c.2098A>G	NM_012433	0.105263	6	51	id5722
SF3B1	2	198266834	SNP	T	C	Missense_Mutation	p.K700E	c.2098A>G	NM_012433	0.071429	5	65	id1002
SF3B1	2	198267359	SNP	C	A	Missense_Mutation	p.K666N	c.1998G>C	NM_012433	0.170543	22	107	id3460
SF3B1	2	198267359	SNP	C	A	Missense_Mutation	p.K666N	c.1998G>C	NM_012433	0.056818	5	83	id16012
SF3B1	2	198267359	SNP	C	G	Missense_Mutation	p.K666N	c.1998G>C	NM_012433	0.078947	3	35	id15991
SF3B1	2	198267359	SNP	C	G	Missense_Mutation	p.K666N	c.1998G>C	NM_012433	0.051282	4	74	id13848

Gene name	Chrom	Start position	Variant Type	Reference Allele	Variant Allele	Variant Classification	Protein Change	cDNA change	Accession	Variant allele fraction	Variant allele count	Reference allele count	ID
SF3B1	2	198267359	SNP	C	G	Missense_Mutation	p.K666N	c.1998G>C	NM_012433	0.174603	11	52	id2084
SF3B1	2	198267359	SNP	C	G	Missense_Mutation	p.K666N	c.1998G>C	NM_012433	0.348485	23	43	id13636
SF3B1	2	198267359	SNP	C	A	Missense_Mutation	p.K666N	c.1998G>C	NM_012433	0.084507	6	65	id9421
SF3B1	2	198267359	SNP	C	A	Missense_Mutation	p.K666N	c.1998G>C	NM_012433	0.061728	5	76	id5804
SF3B1	2	198267359	SNP	C	A	Missense_Mutation	p.K666N	c.1998G>C	NM_012433	0.067669	9	124	id3092
SF3B1	2	198267360	SNP	T	C	Missense_Mutation	p.K666R	c.1997A>G	NM_012433	0.477612	32	35	id12296
SF3B1	2	198267361	SNP	T	C	Missense_Mutation	p.K666E	c.1996A>G	NM_012433	0.097826	9	83	id8760
SF3B1	2	198267361	SNP	T	G	Missense_Mutation	p.K666Q	c.1996A>G	NM_012433	0.321429	27	57	id13638
SF3B1	2	198267371	SNP	G	T	Missense_Mutation	p.H662Q	c.1986C>A	NM_012433	0.055556	5	85	id13868
SF3B1	2	198272802	SNP	G	A	Missense_Mutation	p.R387W	c.1159C>T	NM_012433	0.173469	17	31	id2250
SFRS2	17	74732959	SNP	G	A	Missense_Mutation	p.P95L	c.284C>T	NM_003016	0.27907	12	31	id16191
SFRS2	17	74732959	SNP	G	C	Missense_Mutation	p.P95R	c.284C>G	NM_003016	0.084746	5	54	id16184
SFRS2	17	74732959	SNP	G	T	Missense_Mutation	p.P95H	c.284C>A	NM_003016	0.084746	5	54	id3410
SFRS2	17	74732959	SNP	G	C	Missense_Mutation	p.P95R	c.284C>G	NM_003016	0.228571	8	27	id3332
SFRS2	17	74732959	SNP	G	T	Missense_Mutation	p.P95H	c.284C>A	NM_003016	0.083333	6	66	id3299
SFRS2	17	74732959	SNP	G	A	Missense_Mutation	p.P95L	c.284C>T	NM_003016	0.12766	6	41	id13556
SFRS2	17	74732959	SNP	G	C	Missense_Mutation	p.P95R	c.284C>G	NM_003016	0.196078	10	41	id3026
SFRS2	17	74732959	SNP	G	T	Missense_Mutation	p.P95H	c.284C>A	NM_003016	0.289474	11	27	id15181
SFRS2	17	74732959	SNP	G	T	Missense_Mutation	p.P95H	c.284C>A	NM_003016	0.122807	7	50	id3927
SFRS2	17	74732960	SNP	G	T	Missense_Mutation	p.P95T	c.283C>A	NM_003016	0.087719	5	52	id16192
SFRS2	17	74733113	SNP	A	G	Missense_Mutation	p.Y44H	c.130T>C	NM_003016	0.166667	4	20	id15776
SMC1A	X	53409443	SNP	C	T	Missense_Mutation	p.R1090H	c.3269G>A	NM_006306	0.1	3	27	id8615
SMC3	10	112343707	SNP	C	T	Nonsense_Mutation	p.R360*	c.1078C>T	NM_005445	0.111111	5	40	id6729
STAG1	3	136078015	SNP	G	A	Nonsense_Mutation	p.R971*	c.2911C>T	NM_005862	0.476744	41	45	id16988
STAG2	X	123200061	SNP	T	A	Nonsense_Mutation	p.C711*	c.2133T>A	NM_006603	0.082353	7	78	id3061
STAT3	17	40474420	SNP	C	A	Missense_Mutation	p.D661Y	c.1981G>T	NM_139276	0.075	15	185	id6604
STAT3	17	40474420	SNP	C	A	Missense_Mutation	p.D661Y	c.1981G>T	NM_139276	0.035897	7	188	id13856
STAT3	17	40474420	SNP	C	A	Missense_Mutation	p.D661Y	c.1981G>T	NM_139276	0.0301	9	290	id11824
STAT3	17	40475058	SNP	C	G	Missense_Mutation	p.G618R	c.1852G>C	NM_139276	0.116279	10	76	id15379
SUZ12	17	30315461	DEL	AG	-	Frame_Shift_Del	p.S382fs	c.1146_1147delAG	NM_015355	0.09	7	67	id5830
TBL1XR1	3	176750884	SNP	G	A	Nonsense_Mutation	p.R431*	c.1291C>T	NM_024665	0.054945	5	86	id9755
TET2	10	70442645	SNP	G	A	Missense_Mutation	p.R1656H	c.4967G>A	NM_030625	0.040404	4	95	id15581
TET2	4	106155316	INS	-	A	Frame_Shift_Ins	p.R73fs	c.217_218insA	NM_001127208	0.22	23	80	id16185
TET2	4	106155353	DEL	A	-	Frame_Shift_Del	p.Y85fs	c.254delA	NM_001127208	0.21	13	50	id16184
TET2	4	106155403	DEL	TC	-	Frame_Shift_Del	p.S123fs	c.367_368delTC	NM_001127208	0.05	9	161	id6767
TET2	4	106155544	DEL	G	-	Frame_Shift_Del	p.E170fs	c.508delG	NM_001127208	0.13	12	83	id6962
TET2	4	106155585	DEL	T	-	Frame_Shift_Del	p.D162fs	c.486delT	NM_001127208	0.17	10	49	id14838
TET2	4	106155640	DEL	AT	-	Frame_Shift_Del	p.I381fs	c.541_542delAT	NM_001127208	0.18	12	53	id14877
TET2	4	106155761	INS	-	A	Frame_Shift_Ins	p.T221fs	c.662_663insA	NM_001127208	0.41	29	41	id6024
TET2	4	106155879	DEL	G	-	Frame_Shift_Del	p.L260fs	c.780delG	NM_001127208	0.39	17	27	id15531
TET2	4	106155921	DEL	C	-	Frame_Shift_Del	p.I274fs	c.822delC	NM_001127208	0.15	6	35	id16146
TET2	4	106156030	INS	-	T	Frame_Shift_Ins	p.L311fs	c.931_932insT	NM_001127208	0.29	9	22	id12755
TET2	4	106156120	SNP	C	T	Nonsense_Mutation	p.Q341*	c.1021C>T	NM_001127208	0.297297	11	26	id3816
TET2	4	106156246	SNP	C	T	Nonsense_Mutation	p.Q383*	c.1147C>T	NM_001127208	0.06383	6	88	id13812
TET2	4	106156304	SNP	C	G	Nonsense_Mutation	p.S423*	c.1268C>G	NM_001127208	0.242718	25	78	id12421
TET2	4	106156315	INS	-	T	Frame_Shift_Ins	p.L427fs	c.1279_1280insT	NM_001127208	0.05	10	210	id6395
TET2	4	106156358	DEL	C	-	Frame_Shift_Del	p.S420fs	c.1259delC	NM_001127208	0.06	8	117	id4591
TET2	4	106156371	DEL	C	-	Frame_Shift_Del	p.S424fs	c.1272delC	NM_001127208	0.42	23	32	id16100
TET2	4	106156421	INS	-	T	Frame_Shift_Ins	p.S462fs	c.1385_1386insT	NM_001127208	0.12	11	79	id6669
TET2	4	106156452	DEL	A	-	Frame_Shift_Del	p.I472fs	c.1416delA	NM_001127208	0.41	24	35	id9978
TET2	4	106156540	SNP	C	T	Nonsense_Mutation	p.Q481*	c.1441C>T	NM_001127208	0.105263	6	51	id8613
TET2	4	106156590	INS	-	GTTC	Frame_Shift_Ins	p.T518fs	c.1554_1555insGTTC	NM_001127208	0.07	11	139	id6946
TET2	4	106156625	SNP	C	G	Nonsense_Mutation	p.S530*	c.1589C>G	NM_001127208	0.234043	22	72	id9844
TET2	4	106156665	DEL	T	-	Frame_Shift_Del	p.S543fs	c.1629delT	NM_001127208	0.2	21	83	id8992
TET2	4	106156687	SNP	C	T	Nonsense_Mutation	p.Q530*	c.1588C>T	NM_001127208	0.090909	3	30	id13286
TET2	4	106156693	DEL	T	-	Frame_Shift_Del	p.L532fs	c.1594delT	NM_001127208	0.07	7	91	id5667
TET2	4	106156694	SNP	T	A	Nonsense_Mutation	p.L532*	c.1594delT	NM_001127208	0.137931	4	25	id13788
TET2	4	106156729	SNP	C	T	Nonsense_Mutation	p.R544*	c.1630C>T	NM_001127208	0.089744	7	71	id9033
TET2	4	106156790	SNP	G	A	Nonsense_Mutation	p.W585*	c.1754G>A	NM_001127208	0.066667	7	98	id10707
TET2	4	106156820	DEL	A	-	Frame_Shift_Del	p.Q595fs	c.1784delA	NM_001127208	0.11	12	94	id9756
TET2	4	106156835	INS	-	A	Frame_Shift_Ins	p.L579fs	c.1736_1737insA	NM_001127208	0.04	6	134	id5827
TET2	4	106156862	SNP	C	G	Nonsense_Mutation	p.S588*	c.1763C>G	NM_001127208	0.085714	9	96	id17116
TET2	4	106156862	SNP	C	G	Nonsense_Mutation	p.S588*	c.1763C>G	NM_001127208	0.070423	5	66	id13618
TET2	4	106156936	DEL	G	-	Frame_Shift_Del	p.G634fs	c.1900delG	NM_001127208	0.28	29	73	id6827
TET2	4	106157003	DEL	A	-	Frame_Shift_Del	p.Q656fs	c.1967delA	NM_001127208	0.2	7	28	id9231
TET2	4	106157106	DEL	A	-	Frame_Shift_Del	p.P690fs	c.2070delA	NM_001127208	0.1	7	62	id7339
TET2	4	106157155	DEL	AG	-	Frame_Shift_Del	p.R686fs	c.2056_2057delAG	NM_001127208	0.04	7	160	id5554
TET2	4	106157173	INS	-	A	Frame_Shift_Ins	p.E692fs	c.2074_2075insA	NM_001127208	0.15	18	103	id14876
TET2	4	106157212	SNP	C	T	Nonsense_Mutation	p.Q705*	c.2113C>T	NM_001127208	0.181818	14	63	id3142
TET2	4	106157236	DEL	T	-	Frame_Shift_Del	p.F713fs	c.2137delT	NM_001127208	0.06	6	88	id8785
TET2	4	106157299	SNP	C	T	Nonsense_Mutation	p.Q734*	c.2201A>G	NM_001127208	0.181818	14	63	id12570
TET2	4	106157306	SNP	C	A	Nonsense_Mutation	p.S757*	c.2270C>A	NM_001127208	0.148936	14	80	id10777
TET2	4	106157313	DEL	C	-	Frame_Shift_Del	p.L759fs	c.2277delC	NM_001127208	0.39	36	57	id12296
TET2	4	106157326	DEL	C	-	Frame_Shift_Del	p.Q764fs	c.2290delC	NM_001127208	0.24	26	81	id9503
TET2	4	106157349	DEL	AAAG	-	Frame_Shift_Del	p.I771fs	c.2313_2316delAAAG	NM_001127208	0.05	9	162	id6622
TET2	4	106157371	SNP	C	T	Nonsense_Mutation	p.Q779*	c.2335C>G	NM_001127208	0.068493	5	68	id11567
TET2	4	106157384	INS	-	C	Frame_Shift_Ins	p.H783fs	c.2348_2349insC	NM_001127208	0.08	8	95	id9593
TET2	4	106157407	SNP	C	T	Nonsense_Mutation	p.Q770*	c.2308C>T	NM_001127208	0.087912	8	83	id16702
TET2	4	106157493	INS	-	A	Frame_Shift_Ins	p.E819fs	c.2457_2458insA	NM_001127208	0.16	11	59	id6628
TET2	4	106157615	DEL	A	-	Frame_Shift_Del	p.H839fs	c.2516delA	NM_001127208	0.16	8	43	id5794
TET2	4	106157616	DEL	C	-	Frame_Shift_Del	p.H839fs	c.2516delA	NM_001127208	0.11	11	90	id6042
TET2	4	106157671	DEL	A	-	Frame_Shift_Del	p.H858fs	c.2572delA	NM_001127208	0.18	11	50	id15401
TET2	4	106157695	SNP	C	T	Nonsense_Mutation	p.Q866*	c.2596C>T	NM_001127208	0.207547	11	42	id16170
TET2	4	106157731	DEL	C	-	Frame_Shift_Del	p.L878fs	c.2632delC	NM_001127208	0.29	14	34	id7611
TET2	4	106157816	INS	-	GTCTG	Frame_Shift_Ins	p.M906fs	c.2717_2718insGTCTG	NM_001127208	0.08	6	68	id3168
TET2	4	106157824	SNP	C	T	Nonsense_Mutation	p.Q900*	c.2725C>T	NM_001127208	0.113208	6	47	id15069
TET2	4	106157827	SNP	C	T	Nonsense_Mutation	p.Q910*	c.2728C>T	NM_001127208	0.142857	7	42	id3120
TET2	4	106157833	DEL	GC	-	Frame_Shift_Del	p.A912fs	c.2734_2735delGC	NM_001127208	0.16	9	49	id14836
TET2	4	106157845	SNP	C	T	Nonsense_Mutation	p.Q937*	c.2809C>T	NM_001127208	0.068182	3	41	id9874
TET2	4	106157845	SNP	C	T	Nonsense_Mutation	p.Q916*	c.2746C>T	NM_001127208	0.104478	7	60	id14980
TET2	4	106158065	INS	-	A	Frame_Shift_Ins	p.P989fs	c.2966_2967insA	NM_001127208	0.16	10	53	id8791
TET2	4	106158140	INS	-	A	Frame_Shift_Ins	p.A1014fs	c.3041_3042insA	NM_001127208	0.28	17	44	id2276
TET2	4	106158160	SNP	C	T	Nonsense_Mutation	p.Q1042*	c.3124C>T	NM_001				

Gene name	Chrom	Start position	Variant Type	Reference Allele	Variant Allele	Variant Classification	Protein Change	cDNA change	Accession	Variant allele fraction	Variant allele count	Reference allele count	ID
TET2	4	106158187	SNP	C	T	Nonsense Mutation	p.Q1030*	c.3088C>T	NM_001127208	0.396825	25	38	id2189
TET2	4	106158226	INS	-	AT	Frame_Shift_Ins	p.H1064fs	c.3190_3191insAT	NM_001127208	0.06	6	88	id10415
TET2	4	106158332	DEL	C	-	Frame_Shift_Del	p.T1078fs	c.3233delC	NM_001127208	0.05	6	124	id8806
TET2	4	106158356	DEL	CT	-	Frame_Shift_Del	p.T1107fs	c.3320_3321delCT	NM_001127208	0.08	8	95	id2308
TET2	4	106158408	DEL	TT	-	Frame_Shift_Del	p.N1103fs	c.3309_3310delTT	NM_001127208	0.05	6	105	id14667
TET2	4	106158408	DEL	T	-	Frame_Shift_Del	p.N1103fs	c.3309_3310delTT	NM_001127208	0.22	23	82	id12573
TET2	4	106158439	DEL	AC	-	Frame_Shift_Del	p.T1114fs	c.3340_3341delAC	NM_001127208	0.08	6	69	id15950
TET2	4	106158479	INS	-	T	Frame_Shift_Ins	p.Q1127fs	c.3380_3381insT	NM_001127208	0.33	27	55	id6026
TET2	4	106158503	SNP	G	A	Missense Mutation	p.C1135Y	#N/A	NM_001127208	0.278689	17	44	id9984
TET2	4	106158509	SNP	G	A	Splice_Site	c.e3+1	c.3472_splice	NM_001127208	0.3	27	63	id11641
TET2	4	106158509	SNP	G	A	Splice_Site	c.e3+1	c.3472_splice	NM_001127208	0.153153	17	94	id10803
TNFAIP3	6	138196014	SNP	C	T	Nonsense Mutation	p.Q110*	c.328C>T	NM_006290	0.0625	4	60	id5021
TNFAIP3	6	138196885	SNP	C	T	Nonsense Mutation	p.R183*	c.548G>A	NM_006290	0.09375	9	87	id16445
TNFAIP3	6	138198296	SNP	G	T	Nonsense Mutation	p.E297*	c.889G>T	NM_006290	0.177419	11	51	id15210
TNFRSF14	1	2489781	SNP	G	C	Splice_Site	c.e3-1	c.179_splice	NM_003820	0.102941	7	61	id14548
TNFRSF14	1	2491327	DEL	G	-	Frame_Shift_Del	p.G124fs	c.370delG	NM_003820	0.39	7	11	id12761
TP53	17	7574003	SNP	G	A	Nonsense Mutation	p.R342*	c.1024C>T	NM_001126112	0.067797	4	55	id3959
TP53	17	7577090	SNP	C	T	Missense Mutation	p.R283H	c.848G>A	NM_001126112	0.421053	24	33	id9876
TP53	17	7577090	SNP	C	T	Missense Mutation	p.R283H	c.848G>A	NM_001126112	0.076923	4	48	id9684
TP53	17	7577094	SNP	G	A	Missense Mutation	p.R282W	c.844C>T	NM_001126112	0.071429	3	39	id16017
TP53	17	7577097	SNP	C	T	Missense Mutation	p.D281N	c.841G>A	NM_001126112	0.327586	19	39	id2305
TP53	17	7577100	SNP	T	C	Missense Mutation	p.R280G	c.838A>G	NM_001126112	0.087719	5	52	id16933
TP53	17	7577538	SNP	C	T	Missense Mutation	p.R248Q	c.743G>A	NM_001126112	0.537736	57	49	id4884
TP53	17	7577538	SNP	C	T	Missense Mutation	p.R248Q	c.743G>A	NM_001126112	0.164706	14	71	id4005
TP53	17	7577548	SNP	C	A	Missense Mutation	p.G245C	c.733G>T	NM_001126112	0.056338	8	134	id6889
TP53	17	7577559	SNP	G	C	Missense Mutation	p.S241C	c.722C>G	NM_001126112	0.1875	9	39	id17155
TP53	17	7577578	SNP	T	C	Missense Mutation	p.N235D	c.703A>G	NM_001126112	0.081633	4	45	id14981
TP53	17	7578190	SNP	T	C	Missense Mutation	p.Y220C	c.659A>G	NM_001126112	0.085106	4	43	id5608
TP53	17	7578191	SNP	A	G	Missense Mutation	p.Y220H	c.658T>C	NM_001126112	0.139535	6	37	id3142
TP53	17	7578271	SNP	T	C	Missense Mutation	p.H193R	c.578A>G	NM_001126112	0.0625	3	45	id6883
TP53	17	7578389	SNP	G	A	Missense Mutation	p.R181C	c.541C>T	NM_001126112	0.080645	5	57	id14694
TP53	17	7578400	SNP	G	C	Missense Mutation	p.P177R	c.530C>G	NM_001126112	0.193548	12	50	id5804
TP53	17	7578406	SNP	C	T	Missense Mutation	p.R175H	c.524G>A	NM_001126112	0.047619	4	80	id9682
TP53	17	7578406	SNP	C	T	Missense Mutation	p.R175H	c.524G>A	NM_001126112	0.186047	8	35	id15836
TP53	17	7578406	SNP	C	T	Missense Mutation	p.R175H	c.524G>A	NM_001126112	0.048544	5	98	id9552
TP53	17	7578412	SNP	A	T	Missense Mutation	p.V173E	c.518T>A	NM_001126112	0.089888	8	81	id10986
TP53	17	7578457	SNP	C	T	Missense Mutation	p.R158H	c.473G>A	NM_001126112	0.082474	8	89	id4972
TP53	17	7578463	SNP	C	T	Missense Mutation	p.R156H	c.467G>A	NM_001126112	0.041237	4	93	id13772
TP53	17	7578475	SNP	G	A	Missense Mutation	p.P152L	c.455C>T	NM_001126112	0.127273	7	48	id17168
TP53	17	7578492	SNP	C	T	Nonsense Mutation	p.W146*	c.438G>A	NM_001126112	0.092308	6	59	id13823
TP53	17	7578493	SNP	C	T	Nonsense Mutation	p.W146*	c.438G>A	NM_001126112	0.146341	6	35	id16840
TP53	17	7578536	SNP	T	C	Missense Mutation	p.K132E	c.394A>G	NM_001126112	0.065217	3	43	id15972
TP53	17	7578536	SNP	T	C	Missense Mutation	p.K132E	c.394A>G	NM_001126112	0.219512	9	32	id15950
TP53	17	7579315	INS	-	C	Frame_Shift_Ins	p.C124fs	c.371_372insG	NM_001126112	0.09	13	133	id14877
TP53	17	7579358	SNP	C	A	Missense Mutation	p.R110L	c.329G>T	NM_001126112	0.104762	11	94	id14122
TP53	17	7579359	SNP	G	A	Missense Mutation	p.R110C	c.328C>T	NM_001126112	0.207547	11	42	id16019
TP53	17	7579359	SNP	G	A	Missense Mutation	p.R110C	c.328C>T	NM_001126112	0.439394	29	37	id16285
TP53	17	7579470	INS	-	G	Frame_Shift_Ins	p.P72fs	c.216delC	NM_001126112	0.03	14	418	id7283
TP53	17	7579471	DEL	G	-	Frame_Shift_Del	p.P72fs	c.216delC	NM_001126112	0.17	31	149	id15592
U2AF1	21	44514780	SNP	C	T	Missense Mutation	p.R156H	c.467G>A	NM_006758	0.329897	32	65	id2344
U2AF1	21	44514780	SNP	C	T	Missense Mutation	p.R156H	c.467G>A	NM_006758	0.475	38	42	id6999
U2AF1	21	44524456	SNP	G	A	Missense Mutation	p.S34F	c.101C>T	NM_006758	0.12963	7	47	id11641
U2AF1	21	44524456	SNP	G	A	Missense Mutation	p.S34F	c.101C>T	NM_006758	0.190476	4	17	id4102
U2AF1	21	44527564	SNP	T	C	Missense Mutation	p.D14G	c.41A>G	NM_006758	0.068966	6	81	id15962
ZRSR2	X	15822298	SNP	G	C	Missense Mutation	p.R126P	c.377G>C	NM_005089	0.75	15	5	id8081
ZRSR2	X	1583385	SNP	C	T	Nonsense Mutation	p.R295*	c.883C>T	NM_005089	0.443038	35	44	id13539
ZRSR2	X	15841188	DEL	A	-	Frame_Shift_Del	p.G424fs	c.1272delA	NM_005089	0.58	11	8	id7280

Supplementary Table S4
List of non-hematopoietic genes and variants queried

Gene name	Reported mutations used for variant calling	Accession	Number of variants found
ACVR1B	Frameshift/nonsense, G368spl, R420spl, G219E, G219W, D268N, M345I, I346F, A357T, H358Y, H374R, D376N, A446V, D490N, D490V	NM_004302	0
AKT1	E17K, L52R, Q79K, W80R	NM_005163	0
APC	Frameshift/nonsense c150-1500, R141spl, R216spl, R2204*	NM_001127511	1
APOL2	S127L, S127C, G128D	NM_030882	0
ARHGAP35	Frameshift/ns	NM_004491	0
ARID2	Frameshift/ns, D174Y, D174G, R247H, R247S, R285W, R285Q	NM_152641	0
ATP5B	S386L, K459T	NM_001686	0
B2M	Frameshift/ns	NM_004048	0
BAP1	Frameshift/nonsense, G23spl, D311spl, N78S, R227H, R227C	NM_004656	0
CASP8	Frameshift/ns	NM_033355	4
CDH1	Frameshift/ns, E243K, E243Q, R732Q, P825L	NM_004360	0
CDKN1B	Frameshift/ns	NM_004064	0
CDKN2A	Q50*, V51spl, R58*, G65spl, D74V, D74Y, R80*, H83R, H83Y, D84N, D84V, E88*, D108V, D108N, D108Y, W110*, P114L, P114T, E120*, Y129*, D153spl	NM_000077	0
CHD1	M1I, I1080spl, R1189Q	NM_001270	0
CHD4	E34spl, P251L, P251S, R975H, R1105W, R1105Q, R1162W, R1162Q, R1338I, R877Q, R877W	NM_001273	0
CTNNB1	D32Y, D32A, D32G, D32N, S33C, S33P, S33F, S33Y, G34R, G34E, G34V, I35S, H36Y, S37F, S37C, S37A, T41A, T41I, S45F, S45C, S45Y	NM_001904	0
EGFR	L62R, G63R, R108K, R108G, S220P, R222C, R222L, R252C, R252P, A289V, A289T, H304Y, P596L, P596S, G598V, E709K, G719A, G719D, S768I, S768T, L838M, L858R, L861Q, L861R	NM_005228	0
ERBB2	S310F, S310Y, R678Q, L755S, L755R, D769H, D769N, D769Y, V777L, V842I, R896H, M916I, E1244D, E1244Q	NM_004448	0
ERBB3	M91I, V104L, V104M, D297N, D297Y, D297V, K329E, K329T, T355A, T355I, E928G, S1216F	NM_001982	0
EZH1	Frameshift/ns	NM_001991	0
FGFR3	R248C, S249C, G380R, K650T, K650E, P716H	NM_000142	0
GPS2	Frameshift/ns	NM_004489	1
HRAS	G12A, G12D, G12S, G13R, G13V, G13C, Q61R, Q61L, Q61K	NM_176795	0
KDM5C	Frameshift/ns, E23K, E23Q, E386Q, R585C, R585H, C730F, C730R, E1247K	NM_004187	0
KEAP1	S144F, I145F, V155A, V155F, R169C, R260L, R260Q, R272C, R272H, R320P, R320Q, R320M, R415H, R415C, G417E, G417V, G417R, D422N, G423V, R470S, R470H, R470C, G480W, E493V, E493D, M503I, M503K, W544R, W544C, G603W	NM_012289	0
MAP2K4	Frameshift/nonsense, I73spl, R134W, S184L, S251I, P272spl, L297spl, R287H, P306R, P306H, L362spl, K363spl, L385P, L385V	NM_003010	0
MAP3K1	Frameshift/ns, S1330L, S1330W, C1437S, C1437F	NM_005921	1
MTOR	L1460P, C1483R, C1483Y, C1483F, E1799K, F1888V, F1888I, F1888L, I1973F, T1977K, T1977I, T1977S, V2006L, V2006I, S2215Y	NM_004958	0
NFE2L2	W24C, W24R, Q26R, Q26K, D27G, D27H, D29N, D29H, D29Y, L30F, L30H, G31A, G31V, G31R, R34P, R34G, R34Q, E79K, E79Q, T80P, T80R, T80K, G81V, G81S, G81C, G81D, E82D, A124G, A124V	NM_006164	1
PBRM1	Frameshift/ns	NM_181042	0
PIK3CA	R38C, R38L, R38H, R38S, E39K, E81K, R88Q, C90Y, C90G, R93W, R93Q, G106V, G106R, R108L, R108H, E110del, K111R, K111E, K111N, K111del, L113del, R115L, G118spl, V344G, V344A, V344M, N345K, D350G, D350N, G364R, E365V, E365K, C420R, C420G, G451V, G451R, E453K, E453Q, E453D, P471A, P471L, E542K, E542Q, E545K, E545A, E545Q, E545D, E545G, Q546R, Q546K, E726K, C901F, D939G, M1004I, G1007R, Y1021C, Y1021H, T1025A, T1025S, M1043V, M1043I, N1044K, N1044Y, H1047R, H1047L, G1049R	NM_006218	0
PIK3R1	Frameshift/nonsense, P129L, P129S, R340spl, G376R, Y452C, Y452N, Y452H, D464N, D464Y, R503Q, R514C, R514L, D560H, D560G, D560N, L573P, Y580C, Y580D, M582spl	NM_181523	1
RASA1	Frameshift/nonsense, S122L, S418spl, R789L, R789Q, M802I, P868spl	NM_002890	0
RB1	Frameshift/nonsense, N405spl	NM_000321	0
SEPT12	G56spl, R210spl, I211spl, R242spl, D243spl, R244*	NM_144605	0
SMAD4	Frameshift/ns, R97H, D351H, P356R, P356L, R361H, R361C, G386D, G510E, G510R, D537G, D537Y, D537V, P544S, P544L	NM_005359	0
SMARCA4	P12L, P12S, R381Q, A406T, T786N, T786I, T910M, P913L, P913Q, R966W, R1125G, R1125S, R1125M, R1192C, R1192H, Q1195K, Q1195L, A1231D, A1231T, G1232C, G1232S	NM_003072	0
SPOP	E50K, M117V, S119N, S119R, R121Q, W131G, W131C, F133C, F133L, F133S, D140N, D140H	NM_001007228	0
STK11	Frameshift/nonsensesplice-site	NM_000455	0
VHL	Frameshift/nonsensesplice-site, N78S, N78D, N78Y, W88R, W88L, L89P, L89H, S111R, S111H, H115N, V130L, V130D, I151N, I151T, I151F, L158P, L158V, C162R, C162F, C162Y, L169P, L184P, L188P, L188R	NM_000551	1
Total			10

Supplementary Table S5
Called variants in non-hematopoietic genes

Gene name	Chromosome	Start position	Variant Type	Reference Allele	Variant Allele	Variant Classification	Protein Change	cDNA Change	Accession	Variant allele fraction	Variant allele count	Reference allele count	ID
APC	5	112175212	DEL	AAAAG	-	Frame_Shift_Del	p.I1307fs	c.3921_39;NM_001127511		0.28	29	74	id3447
CASP8	2	202149654	DEL	C	-	Frame_Shift_Del	p.N306fs	c.918delC	NM_033355	0.51	94	89	id8185
CASP8	2	202149654	DEL	C	-	Frame_Shift_Del	p.N306fs	c.918delC	NM_033355	0.47	95	109	id7570
CASP8	2	202149901	SNP	C	T	Nonsense_Mutation	p.Q389*	c.1165C>T	NM_033355	0.387755	19	30	id16144
CASP8	2	202149901	SNP	C	T	Nonsense_Mutation	p.Q389*	c.1165C>T	NM_033355	0.385714	27	43	id4204
GPS2	17	7217689	SNP	G	G	Nonsense_Mutation	p.Q80*	c.238C>T	NM_004489	0.098361	12	110	id11021
MAP3K1	5	56152535	SNP	G	G	Nonsense_Mutation	p.W197*	c.591G>A	NM_005921	0.422414	49	67	id6400
NFE2L2	2	178098009	SNP	G	A	Missense_Mutation	p.A124V	c.371C>T	NM_006164	0.069767	3	40	id14589
PIK3R1	5	67589168	SNP	C	C	Nonsense_Mutation	p.R386*	c.1156C>T	NM_181523	0.107143	3	25	id11720
VHL	3	10191614	SNP	C	T	Nonsense_Mutation	p.Q203*	c.607C>T	NM_000551	0.333333	13	26	id13940

Supplementary Table S6
Logistic regression for factors associated with mutations

Logistic regression was performed using the variables age (as a continuous variable), ancestry, sex, T2D, and age/sex interaction. Other interaction terms were modeled, but none were significant. Proportion of variance explained is derived by analysis of variance (ANOVA) for the generalized linear model, and is equal to deviance for the variable divided by residual deviance for the null model.

	Beta coefficient	OR(95 CI)	p-value	Variance explained
Age	0.07	1.08(1.07-1.09)	<0.001	0.06
European (referent)				0.003
African-American	0.15	1.16(0.93-1.44)	0.19	
East-Asian	-0.08	0.92(0.71-1.2)	0.56	
Hispanic	-0.37	0.69(0.56-0.85)	<0.001	
South Asian	-0.25	0.78(0.6-1)	0.057	
No T2D (referent)				0.002
Has T2D	0.28	1.32(1.14-1.54)	<0.001	
Male (referent)				0.001
Female	1.01		0.026	
BMI	-0.02	0.98(0.96-0.99)	0.005	0.001
Age:Female	-0.02		0.009	0.001

Supplementary Table S7**Logistic regression for factors associated with mutations by ancestry group**

Logistic regression was performed using the variables age (as a continuous variable), sex, T2D, and age/sex interaction for each ancestry group. Proportion of variance explained is derived by analysis of variance (ANOVA) for the generalized linear model, and is equal to deviance for the variable divided by residual deviance for the null model.

African-American

	Beta coefficient	OR(95 CI)	p-value	Variance explained
Age	0.07	1.08(1.05-1.1)	<0.001	0.08
Female	0.37	1.45(0.21-10.2)	0.7	0.002
Has T2D	0.24	1.27(0.9-1.79)	0.18	0.002
Age:Sex	-0.01	0.99(0.96-1.02)	0.52	0.0003

East Asian

	Beta coefficient	OR(95 CI)	p-value	Variance explained
Age	0.1	1.11(1.06-1.16)	<0.001	0.03
Female	1.3	3.65(0.08-172)	0.51	0.003
Has T2D	-0.12	0.89(0.57-1.38)	0.61	0.0003
Age:Sex	-0.03	0.97(0.92-1.03)	0.385	0.0006

European

	Beta coefficient	OR(95 CI)	p-value	Variance explained
Age	0.07	1.07(1.06-1.09)	<0.001	0.05
Female	1.73	5.62(1.26-25)	0.023	0.001
Has T2D	0.31	1.36(1.04-1.79)	0.026	0.003
Age:Sex	-0.03	0.97(0.95-0.99)	0.013	0.0033

Hispanic

	Beta coefficient	OR(95 CI)	p-value	Variance explained
Age	0.08	1.08(1.06-1.11)	<0.001	0.08
Female	0.69	2(0.27-15.6)	0.5	0.00001
Has T2D	0.22	1.24(0.9-1.72)	0.18	0.001
Age:Sex	-0.01	0.99(0.96-1.02)	0.485	0.0003

South Asian

	Beta coefficient	OR(95 CI)	p-value	Variance explained
Age	0.08	1.08(1.05-1.11)	<0.001	0.07
Female	-0.82	0.44(0.01-12.3)	0.64	0.0001
Has T2D	0.49	1.63(1.04-2.56)	0.033	0.007
Age:Sex	0.01	1.01(0.96-1.07)	0.677	0.0002

Supplementary Table S8

Details on subjects that developed hematologic malignancies

Incident Cases

Age at sampling	Diagnosis	Cohort	Adjudicated	Latency (years)	Mutation on WES (VAF)	Mutations on RHP (VAF)	WBC	HGB	PLT	Death	Cause of death
77	CANCER OF SPLEEN ^a	AJ	No	6	JAK2 p.V617F (0.23)	NA	7.8	11	247	Yes	CARCINOMA UNSPECIFIED SITE
64	LEUKEMIA (prior NHL) ^b	AJ	No	7	ASXL1 p.D616fs (0.23)	ASXL1 p.D616fs (0.18)	3.5	12.9	189	No	
57	LYMPHOMA ^c	AJ	No	2	DNMT3A p.R882H (0.29)	NA	14.3 (51.3% lymphocytes) ^d	11	248	No	
85	DLBCL, large intestine ^e	MEC	Yes	5	TET2 p.C1135Y (0.28)/ASXL1 p.I919fs (0.22)	TET2 p.C1135Y (0.35)/ASXL1 p.I919fs (0.26)/TET2 p.G1192V (0.30) ^f				No	
82	MDS-RAEB	MEC	Yes	7	ASXL1 p.T514fs (0.24)	ASXL1 p.T514fs (0.30)/TET2 p.1616fs (0.15) ^f				Yes	Myeloid leukemia
50	LYMPHOMA	AJ	No	4	None	NA	6.3	14.4	220	Yes	COMPLICATIONS DUE TO LYMPHOMA
48	LYMPHOMA	AJ	No	1	None	NA	4.6	14.3	225	Yes	RENAL FAILURE
57	LYMPHOMA	AJ	No	7	None	None	7.4	13.7	280	Yes	HYPOXIC RESPIRATORY FAILURE
62	LEUKEMIA	AJ	No	9	None	NA	2.6	13.9	168	No	
51	BLOOD	AJ	No	8	None	NA	5.7	13.7	318	No	
67	LEUKEMIA	AJ	No	9	None	None				No	
64	MULTIPLE MYELOMA	AJ	No	9	None	NA	4.4	11.4	298	No	
59	LYMPHOMA	AJ	No	8	None	NA		13.2	195	No	
61	ACUTE MYELOID LEUKEMIA	AJ	No	10	None	None		11.7	158	Yes	ACUTE MYELOID LEUKEMIA
51	LEUKEMIA	AJ	No	10	None	NA	4.8	11	334	No	
66	Multiple myeloma	MEC	Yes	8	None	NA				No	

Prevalent Cases

Age at sampling	Diagnosis	Cohort	Adjudicated	Time (years prior)	Mutation on WES (VAF)	Mutations on RHP (VAF)	WBC	HGB	PLT	Death	Cause of death
64	LEUKEMIA	AJ	No	7	None	NA	3.5	12.9	189	No	
48	LYMPHOMA	AJ	No	1	None	NA	4.6	14.3	225	Yes	RENAL FAILURE
76	LEUKEMIA	AJ	No	12	None	NA		11.4	252	No	
63	NHL oropharynx	MEC	Yes	24	None	NA					
64	NHL (subsequent LEUKEMIA, see above)	AJ	No	7	ASXL1 p.D616fs (0.23)	ASXL1 p.D616fs (0.18)	3.5	12.9	189	No	

NHL=Non-Hodgkin's lymphoma, DLBCL=diffuse large B-cell lymphoma, MDS-RAEB=myelodysplastic syndrome, refractory anemia with excess blasts

AJ= Jackson Heart Study, MEC= Multi-ethnic cohort, Hispanics in Los Angeles

WES=whole exome sequencing, RHP=Rapid Heme Panel targeted re-sequencing, VAF=variant allele fraction

WBC=white blood cell count (x 10⁹ cells/L), HGB=hemoglobin (g/dL), PLT=platelet count (x10⁹cells/L)

NA=Not available

a This is likely splenomegaly secondary to a JAK2-mutated myeloproliferative neoplasm

b This likely represents a therapy related AML/MDS, as ASXL1 mutations have never been described in lymphoid malignancies

c DNMT3A mutations have been found in peripheral T-cell lymphomas and early T-progenitor ALL

d This is the second highest absolute lymphocyte count in the cohort of 2,426 subjects who had a WBC differential. The subject with the highest absolute lymphocyte count also had a DNMT3A mutation.

e TET2 has been reported to be mutated in 6-12% of DLBCL, and the mutations are reported to be found in hematopoietic stem cells

f Mutations outside of exon 3 of TET2 were not detectable by WES, but were detected by RHP

Supplementary Table S9
Logistic regression for factors associated with RDW \geq 14.5%

Only individuals with an MCV>80 fL were included in this analysis. Individuals were from Jackson Heart Study or UA control cohort.

Covariate	OR (95% CI)	p-value
Age 60-69	1.3(0.9-1.8)	0.17
Age 70-79	1.8(1.2-2.7)	0.002
Age 80-89	2.7(1.3-5.2)	0.005
Age >90	1.7(0.9-3.1)	0.09
Female	1.1(0.8-1.5)	0.76
Has T2D	0.9(0.7-1.2)	0.45
DNMT3a mutation	1.9(1-3.6)	0.048
WBC	1.1(1.0-1.1)	0.092
Hemoglobin	0.7(0.7-0.8)	<0.001
Platelet count	1.0(1.0-1.0)	0.65

Supplementary TableS10 Association of cytopenias with mutations

Individuals were classified as having cytopenia as defined in Methods. Statistical comparisons were performed using Fisher's exact test. WBC- white blood cell count, Hgb - hemoglobin, Plt - platelet count. Individuals were from Jackson Heart Study, Longevity Genes Project, Botnia, Helsinki-sib, and Malmo-sib. Of the 5 individuals with multiple cytopenias and clonal mutations, 2 had 2 detectable mutations, suggesting that these might be undiagnosed cases of MDS.

	Mutation	No Mutation	
Low WBC	6	59	65
Normal WBC	119	2628	2747
	125	2687	

OR 2.2(0.8-5.2), p=0.066

	Mutation	No Mutation	
Low Hgb	30	616	646
Normal Hgb	109	2350	2459
	139	2966	

OR 1.0(0.7-1.6), p=0.83`

	Mutation	No Mutation	
Low Plt	4	107	111
Normal Plt	132	2789	2921
	136	2896	

OR 0.8(0.2-2.1), p=0.82

	Mutation	No Mutation	
Any cytopenia	35	745	780
No cytopenia	104	2224	2328
	139	2969	

OR 1.0(0.6-1.5), p=1

	Mutation	No Mutation	
2+ cytopenias 1 or 0 cytopenias	5	37	42
	134	2932	3066
	139	2969	

OR 3.0(0.9-7.7), p=0.037

Lower limit of normal for blood counts were defined as follows:

White blood cell count:

African-Americans $3.0 \times 10^9/L$, white $4.0 \times 10^9/L$

Hemoglobin:

African-American women 11.2 g/dL,

African-American men 12.5 g/dL, white women

11.9 g/dL, white men 13.4 g/dL

Platelet count:

$150 \times 10^9/L$

Supplementary Table S11

Association of mutations with known versus unknown causes of anemia

During clinical evaluation, most patients with anemia can be found to have an attributable cause. Using subjects from the Jackson Heart Study, we assessed whether the anemia was attributable to iron deficiency, anemia of chronic inflammation, or renal insufficiency. Individuals with clonality were less likely to have anemia attributable to one of these causes.

Iron deficiency anemia/microcytic anemia

- mean corpuscular volume < 80 fL
- ferritin < 20 ng/mL
- ferritin 20-100 ng/mL WITH EITHER total iron binding capacity > 370 mcg/dL OR serum iron < 50 mcg/dL OR iron saturation < 20%

Anemia of chronic disease

- serum iron < 65 mcg/dL WITH total iron binding capacity < 250 mcg/dL
- ferritin > 350 ng/mL for males
- ferritin > 300 ng/mL for females

Renal insufficiency

- estimated glomerular filtration rate < 30 mL/min/1.73m²

	Mutation	No mutation	
Known cause	2	360	362
Unknown cause	7	151	158
	9	511	

OR 0.1(0-0.6), p=0.004

Supplementary Table S12

Risks associated with developing incident coronary heart disease and ischemic stroke using traditional risk factors and mutations

Hazard ratios were estimated using competing risks regression with death as the competing risk. P-values are derived from the Fine-Gray test. Individuals with prior coronary heart disease (CHD) were excluded for CHD analysis, and individuals with prior ischemic stroke were excluded for stroke analysis. Models shown in A or B are from the same population and differ by having covariates either removed or added. A) Coronary heart disease, B) ischemic stroke. Individuals were from Jackson Heart Study and FUSION.

A

Covariate	Model 1		Model 2		Model 3	
	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
log(Age)	5.3(2.1-12.8)	<0.001	4.7(1.9-11)	<0.001	4.6(1.9-11)	<0.001
Has T2D	3.3(2.1-5.3)	<0.001	3.4(2.1-5.4)	<0.001	3.5(2.2-5.5)	<0.001
Female	0.7(0.5-1.1)	0.11	0.7(0.5-1.1)	0.13	0.7(0.5-1.1)	0.13
HDL<35 mg/dL	1.1(0.6-2.1)	0.77	1.1(0.6-2.1)	0.81	1.1(0.6-2.1)	0.78
HDL>60 mg/dL	0.7(0.4-1.2)	0.18	0.7(0.4-1.3)	0.25	0.7(0.4-1.3)	0.26
TC >240 mg/dL	2.1(1.3-3.2)	<0.001	2(1.3-3.1)	<0.001	2(1.3-3.1)	<0.001
Former or current smoker	1.6(1.1-2.5)	0.024	1.6(1-2.4)	0.035	1.6(1.1-2.5)	0.02
Hypertension stage II-IV	1.6(1-2.5)	0.06	1.4(0.9-2.3)	0.15	1.4(0.9-2.3)	0.15
BMI>25	1.2(0.6-2.5)	0.55	1.4(0.6-2.8)	0.43	1.3(0.6-2.8)	0.42
Mutation present			2.3(1.1-4.8)	0.026		
VAF<0.10					1.4(0.5-4)	0.55
VAF≥0.10					4.4(1.9-10.5)	<0.001
Pseudo Log-likelihood		-661		-658		-656
Pseudo likelihood ratio test		103 on 9 df		109 on 10 df		113 on 11 df

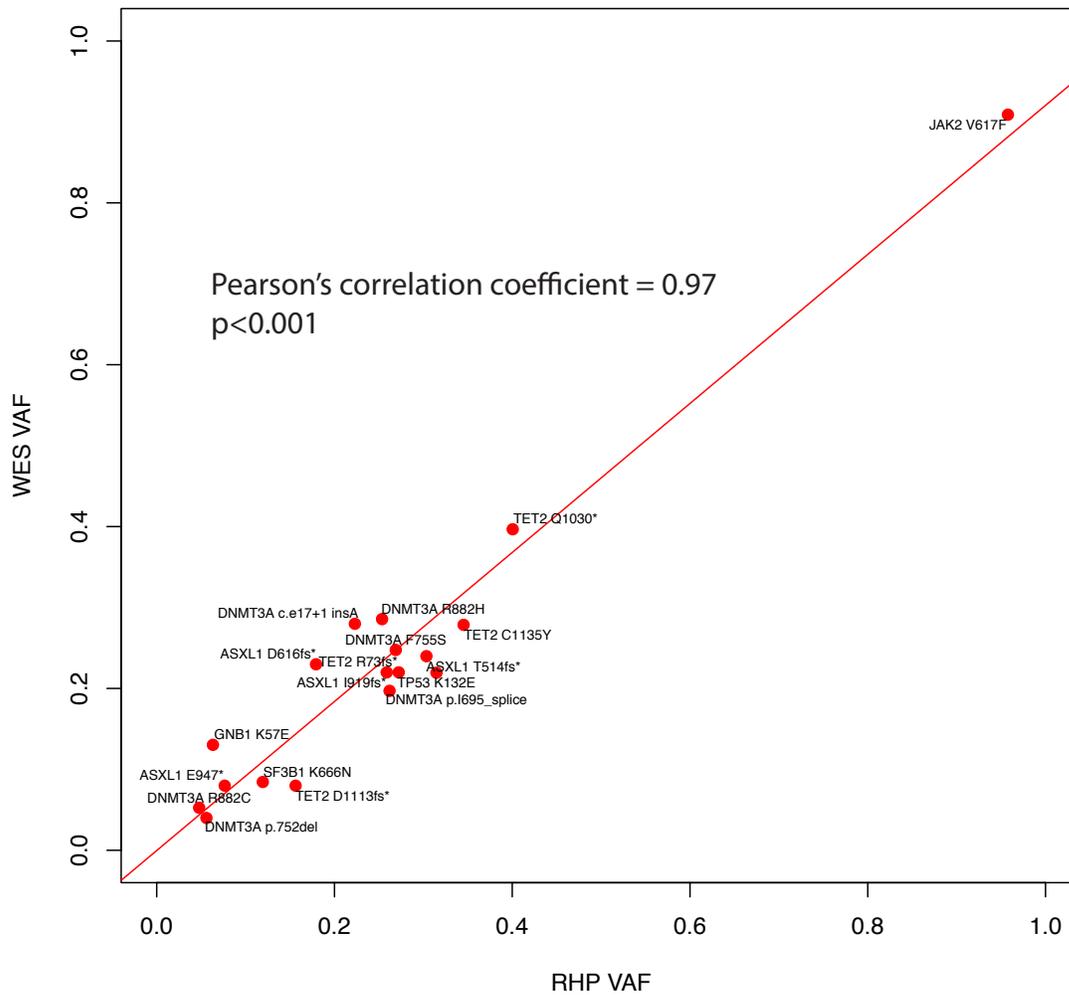
B

Covariate	Model 1		Model 2		Model 3	
	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
log(Age)	14.6(4.8-44.7)	<0.001	13.3(4.3-40)	<0.001	13.1(4.3-40)	<0.001
Has T2D	2.9(1.7-5)	<0.001	2.9(1.7-5)	<0.001	3(1.7-5.2)	<0.001
Female	0.8(0.5-1.3)	0.44	0.9(0.5-1.4)	0.53	0.9(0.5-1.4)	0.55
HDL<35 mg/dL	1.2(0.6-2.4)	0.52	1.3(0.7-2.5)	0.45	1.3(0.7-2.5)	0.45
HDL>60 mg/dL	1(0.6-1.9)	0.95	1.1(0.6-2)	0.84	1.1(0.6-2)	0.86
TC >240 mg/dL	1.3(0.8-2.1)	0.29	1.3(0.8-2.1)	0.31	1.3(0.8-2.1)	0.29
Former or current smoker	1.8(1.1-2.9)	0.014	1.8(1.1-2.9)	0.016	1.8(1.1-2.9)	0.014
Hypertension stage II-IV	1.8(1-3.1)	0.037	1.7(0.9-2.9)	0.077	1.7(1-2.9)	0.074
BMI>25	1.5(0.6-3.6)	0.35	1.6(0.7-4)	0.3	1.6(0.7-3.9)	0.3
Mutation present			2.2(1.1-4.6)	0.029		
VAF<0.10					1.8(0.7-4.6)	0.2
VAF≥0.10					3.1(1.2-8.4)	0.025
Pseudo Log-likelihood		-464		-462		-462
Pseudo likelihood ratio test		79 on 9 df		83 on 10 df		83.5 on 11 df

Supplementary Figure S1

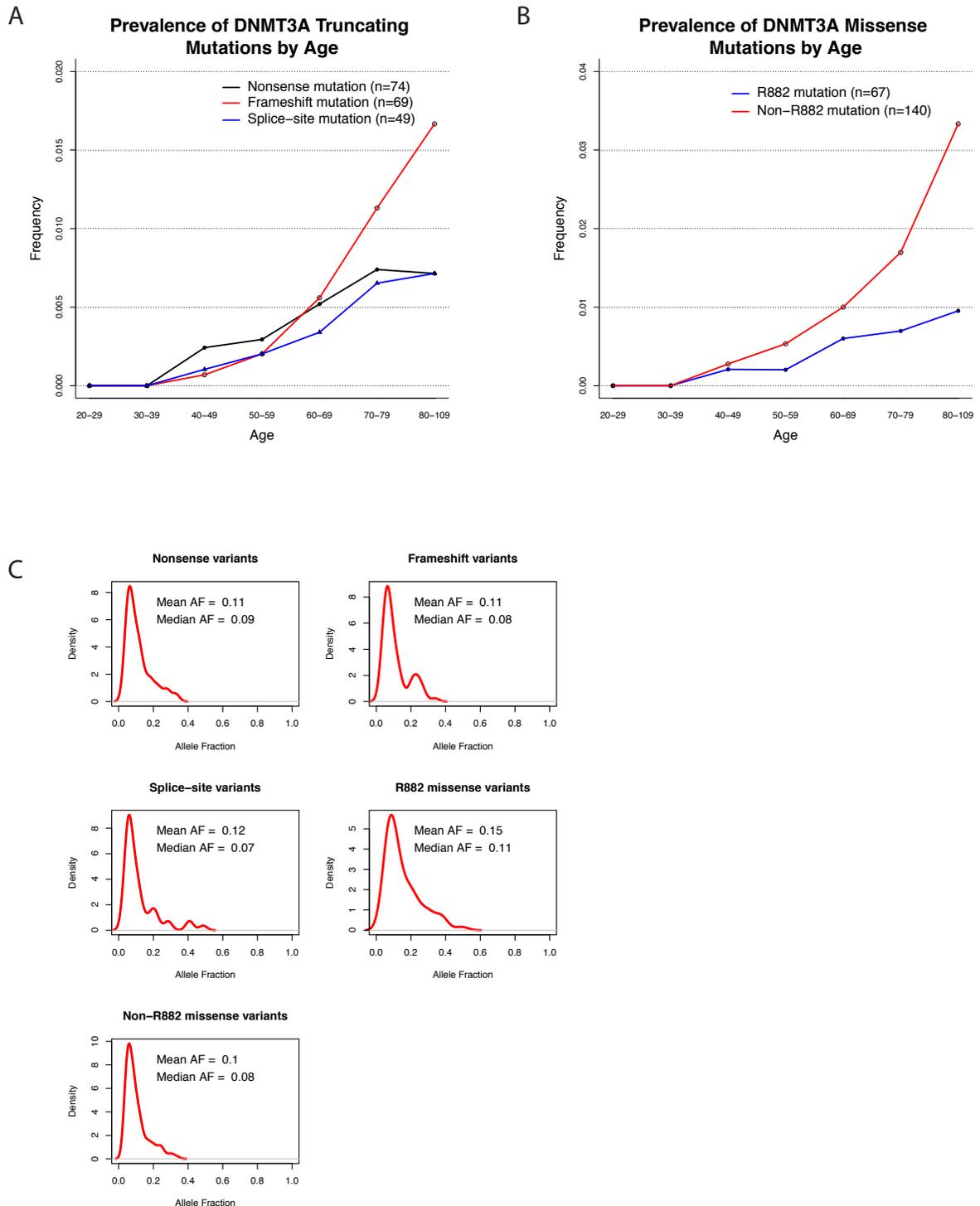
Validation of putative somatic variants

Eighteen variants were validated using targeted, amplicon based re-sequencing ("Rapid Heme Panel", see Methods). Comparison of variant allele fraction (VAF) between the two methods is shown for the 18 variants. WES, whole exome sequencing. RHP, Rapid Heme Panel.



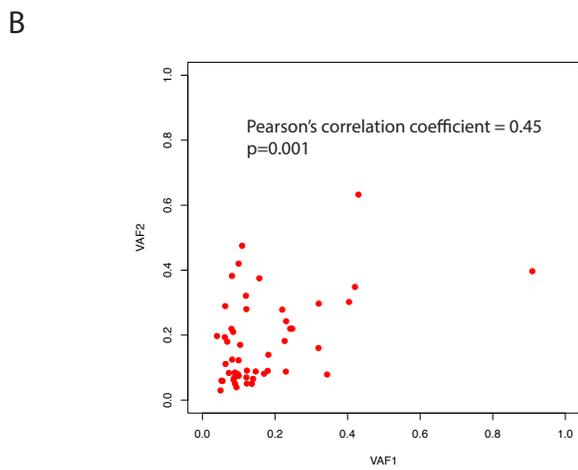
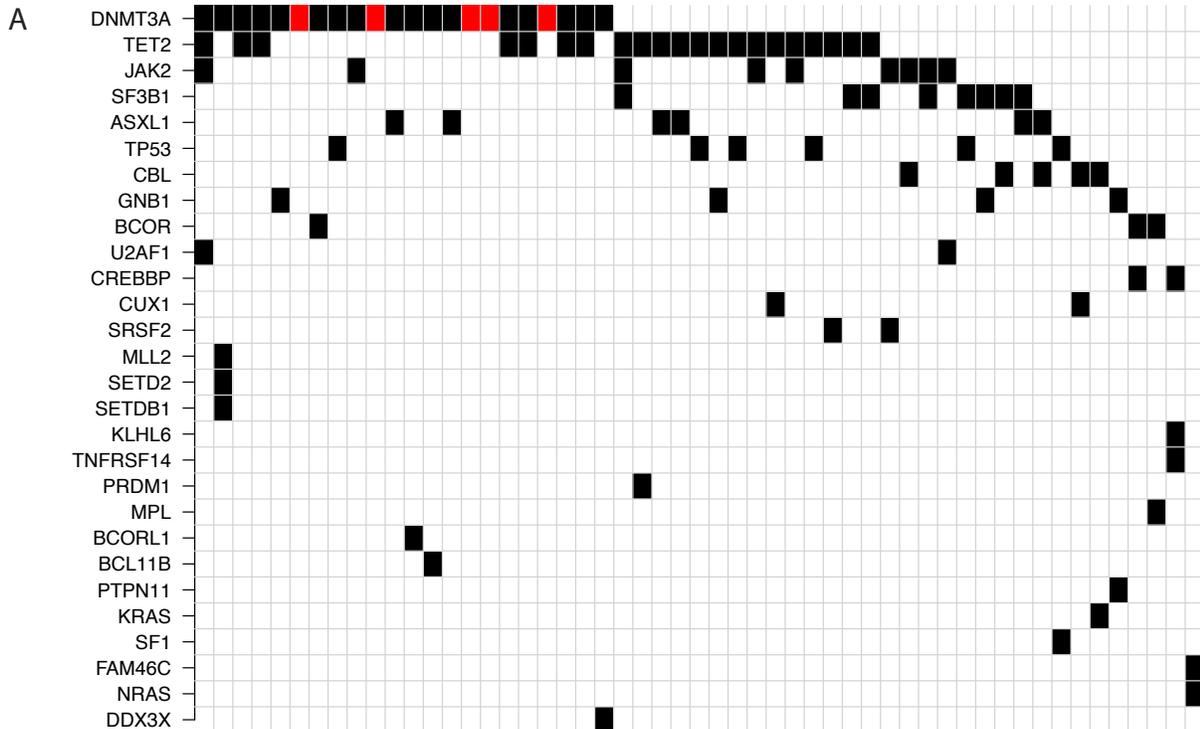
Supplementary Figure S2 Characteristics of *DNMT3A* variants

A) Frequency of *DNMT3A* nonsense, frameshift, and splice-site variants called as somatic by age group. B) Frequency of R882 and non-R882 missense variants called as somatic by age group. C) Allele fraction of called *DNMT3A* variants by mutation type.



Supplementary Figure S3 Co-mutations

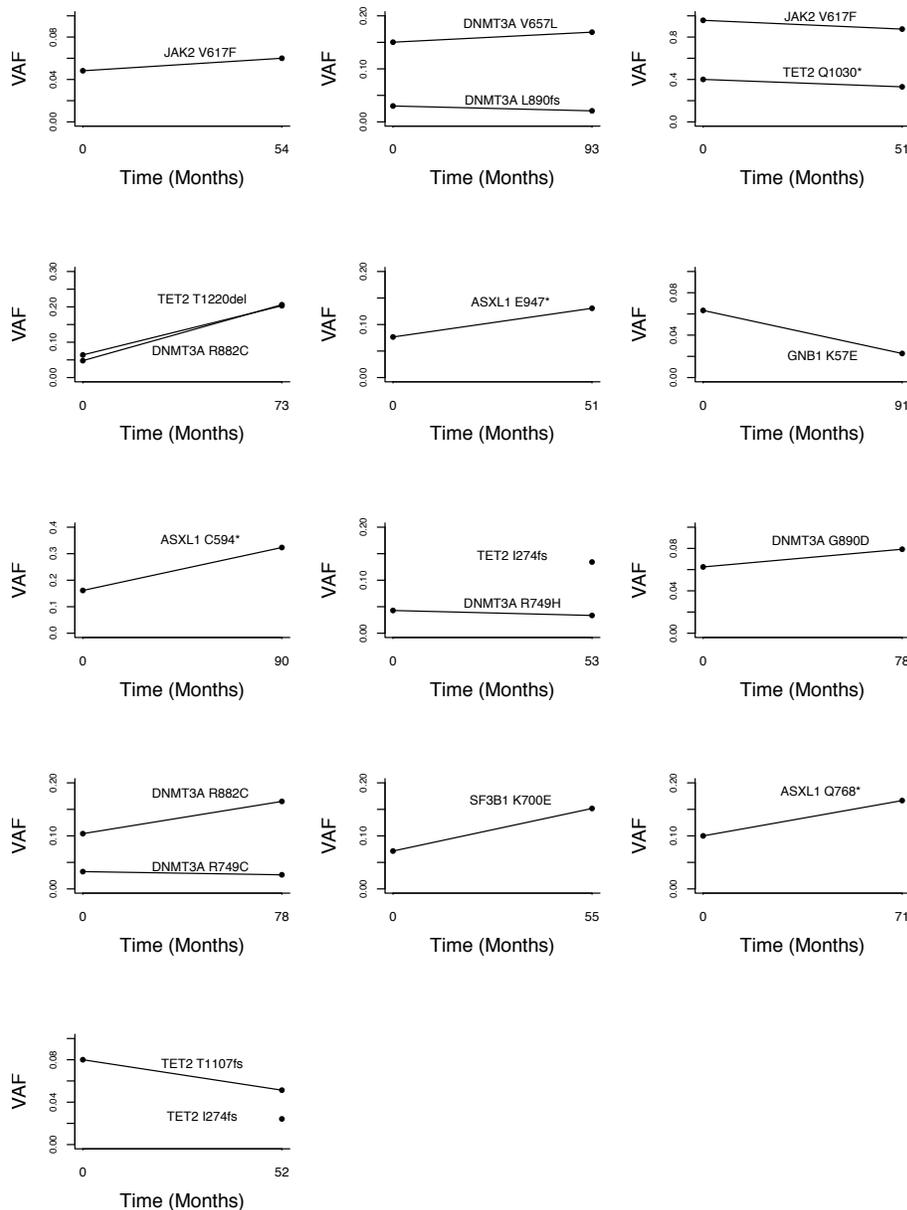
A) Co-mutation plot, individuals are represented by columns. Black rectangles represent mutated genes, red rectangles represent 2 separate mutations in the same gene.
B) Correlation plot for variant allele fraction (VAF) from the 49 subjects with 2 mutations.



Supplementary Figure S4

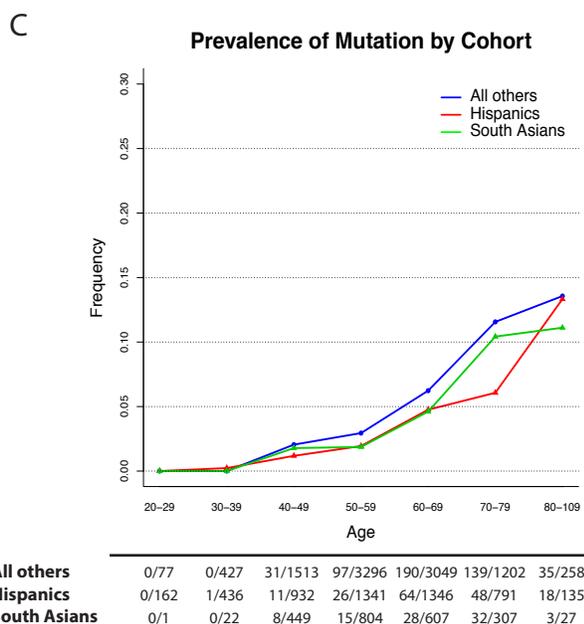
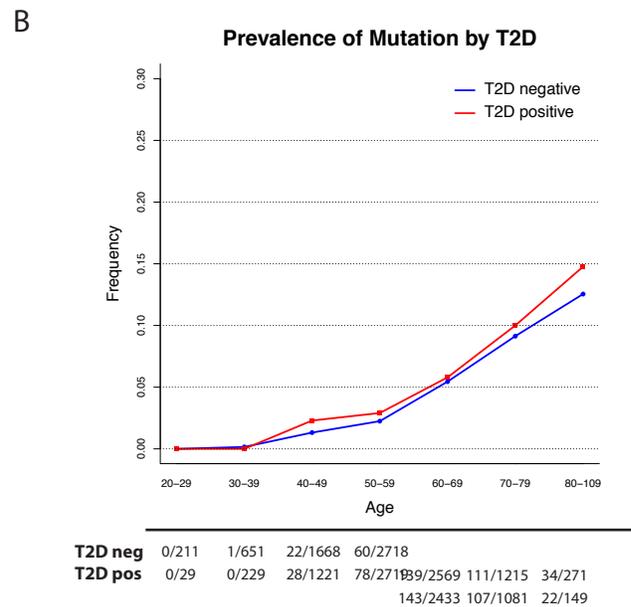
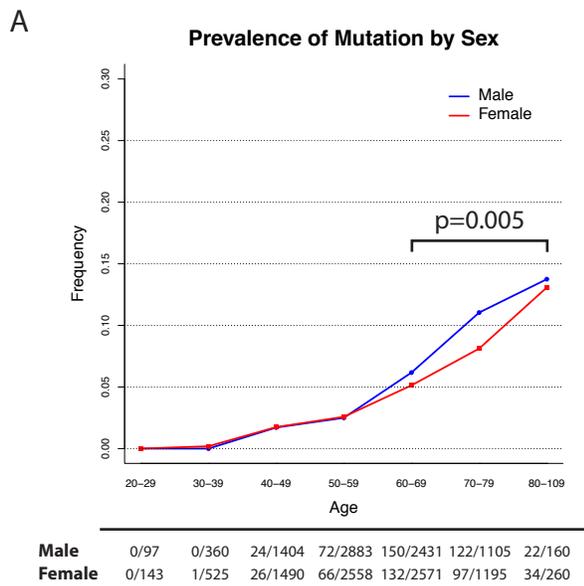
Longitudinal assessment of mutations

Peripheral blood DNA from 4 to 8 years after the initial DNA collection was available for 13 subjects in JHS who had detectable mutations on exome sequencing. As described in the Supplementary Methods, amplicon-based targeted re-sequencing was performed for a panel of 95 genes. Graphs represent individual subjects, with mutation variant allele frequency (VAF) from the initial and later time points shown. All of the initial mutations detected on exome sequencing were still present in the later sample, and 2 subjects had acquired new mutations.



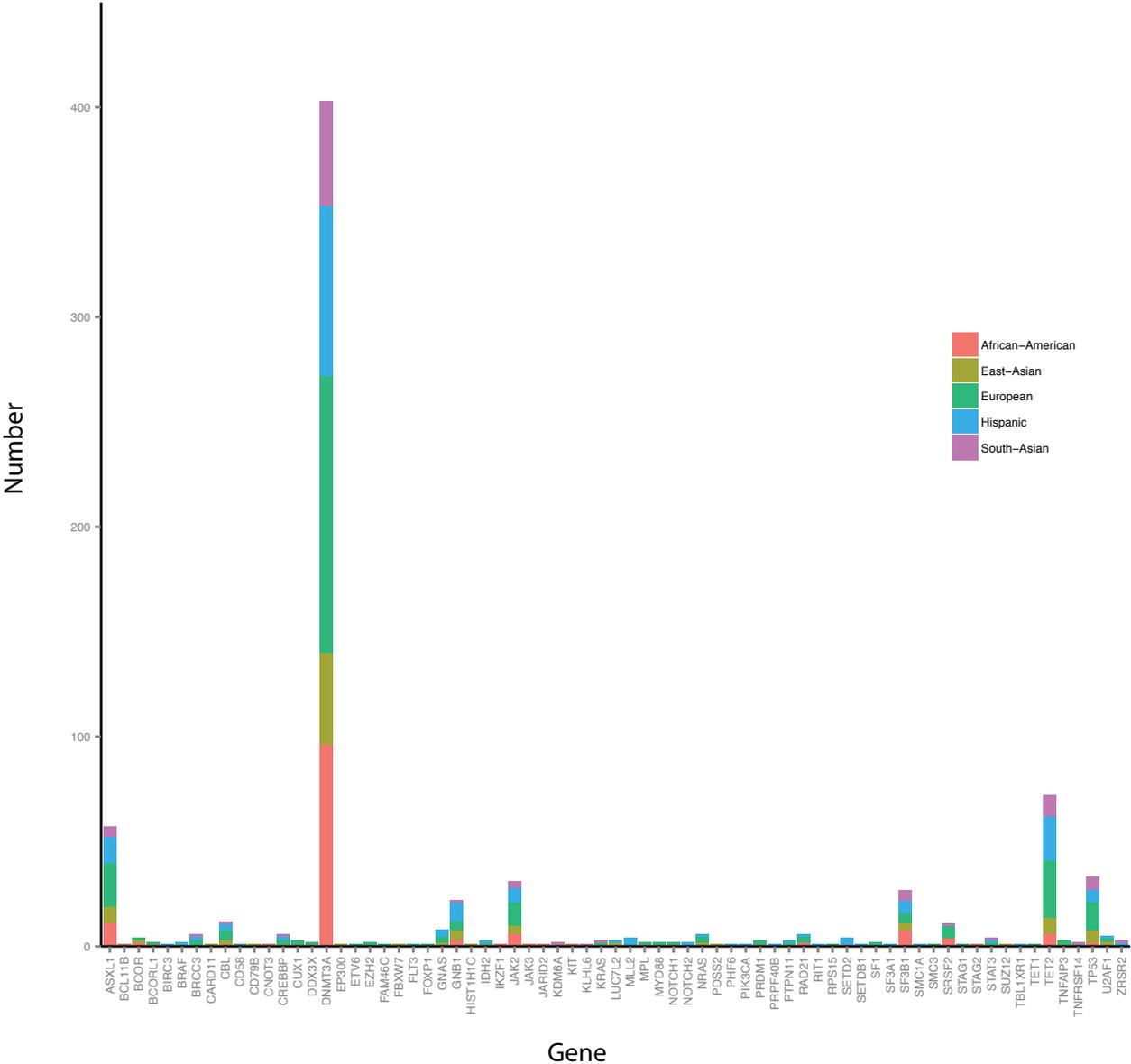
Supplementary Figure S5 Factors associated with mutations

A) Frequency of mutation for males and females by age group. For those 60 or older, being male is associated with having a detectable clone (OR 1.3, 95% CI 1.1-1.5, $p=0.005$ by multivariable logistic regression using age, sex, T2D and BMI as covariates). B) Frequency of mutation for those with and without type 2 diabetes by age group. C) Frequency of mutation for non-Hispanics, Hispanics, and South Asians by age group.



Supplementary Figure S6 Mutations by ethnic background

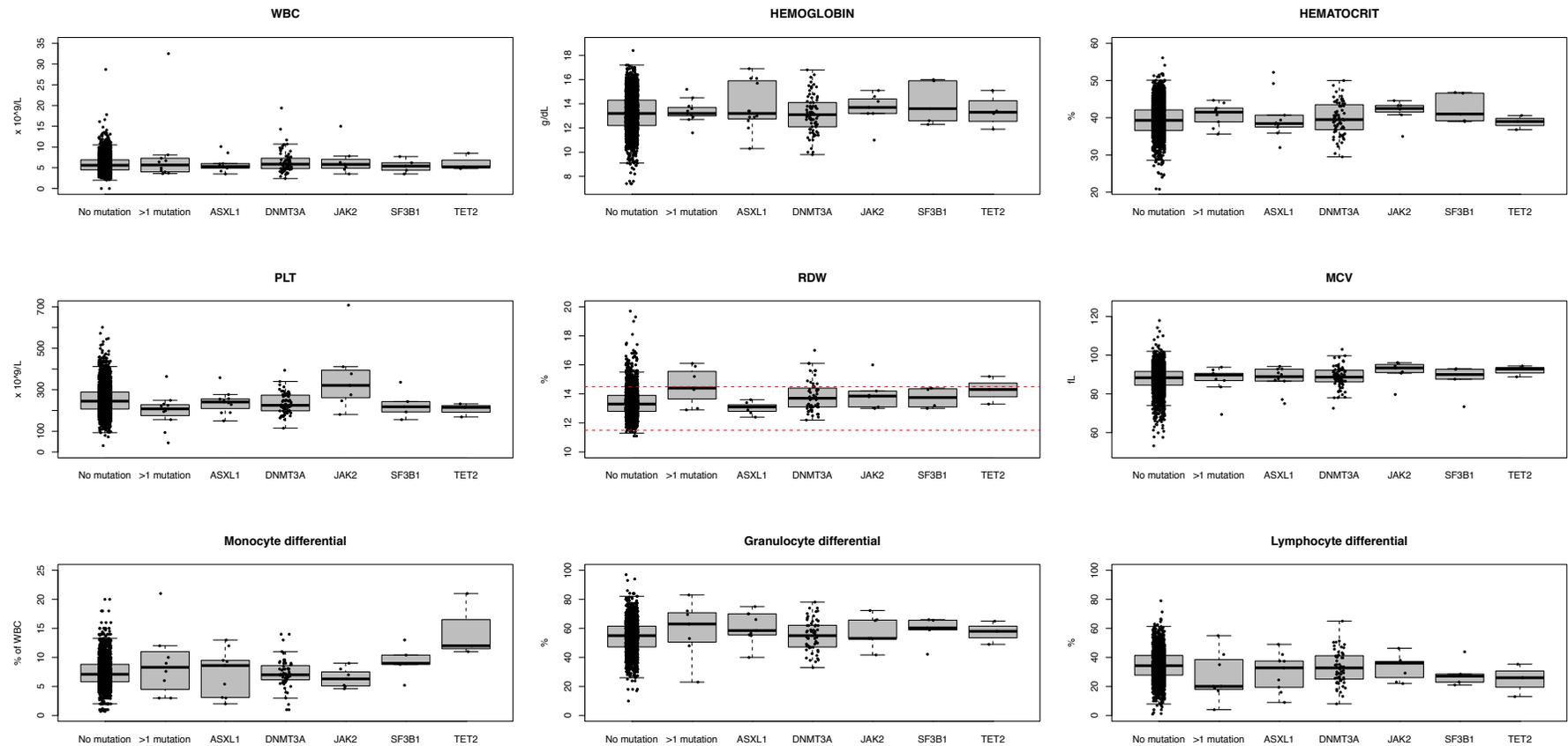
Number of mutations for each gene stratified by ethnic background.



Supplementary Figure S7

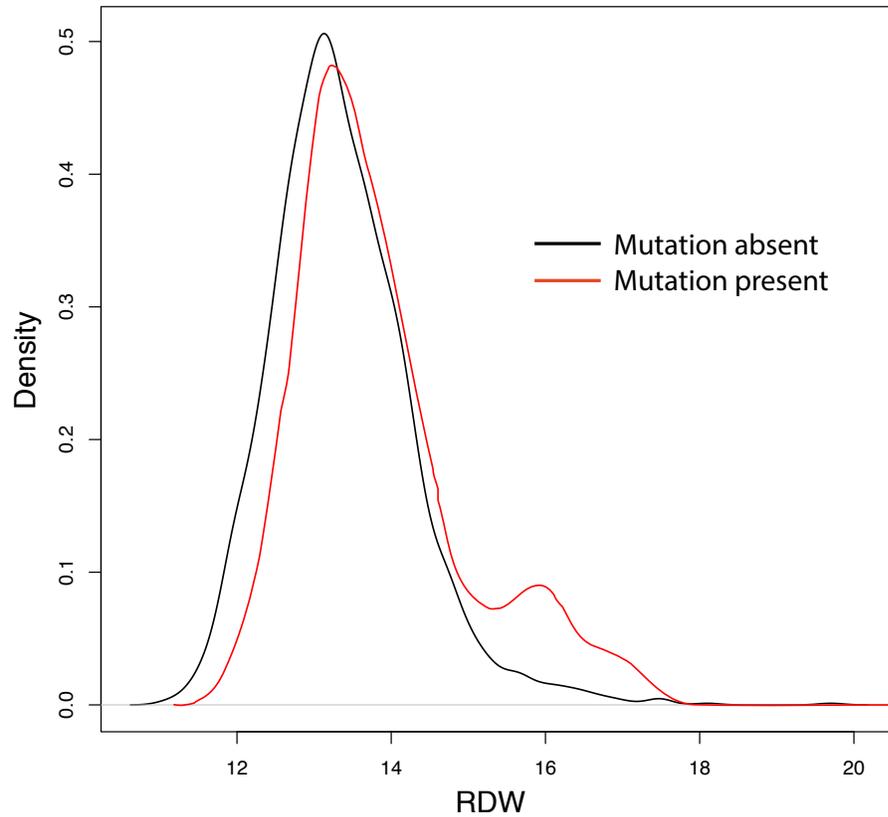
Blood counts for individuals with and without detectable mutations

Dots represent individuals. Box represents 25th and 75th percentiles, line in box represents median. Whiskers represent 5th and 95th percentiles. For listed genes, individuals only had mutations in that gene, and not other genes. Dashed red lines represent 11.5% and 14.5%, the normal ranges for RDW. Abbreviations: WBC-white blood cell count, PLT-platelet count, RDW-red cell distribution width, MCV-mean corpuscular volume. Individuals were from Jackson Heart Study, Longevity Genes Project, Botnia, Helsinki-sib, or Malmo-sib.

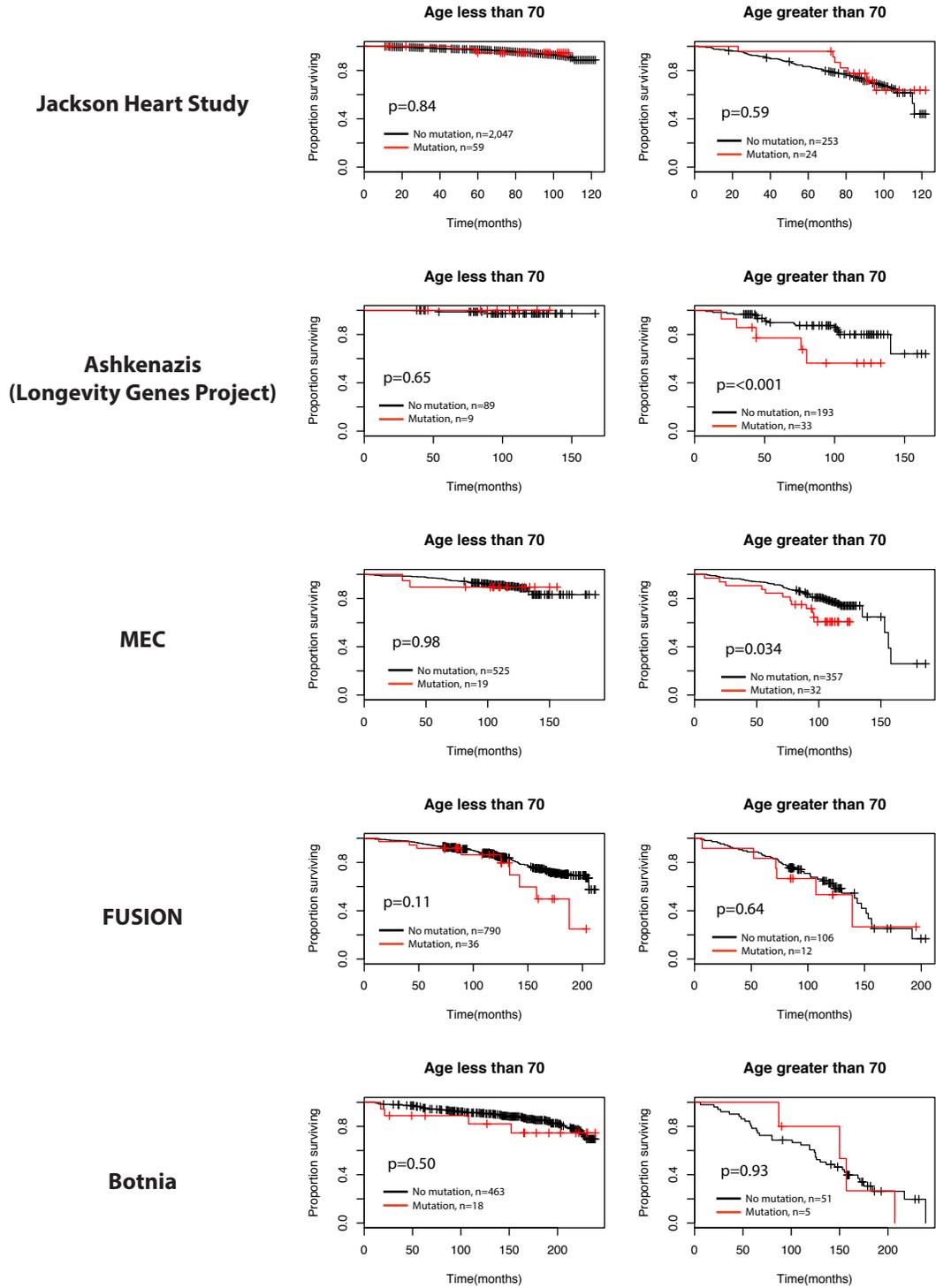


Supplementary FigureS8
Density plot for RDW

RDW in subjects with and without mutations. Only subjects with an MCV>85 fL were included in this analysis. Individuals were from Jackson Heart Study or Longevity Genes Project.

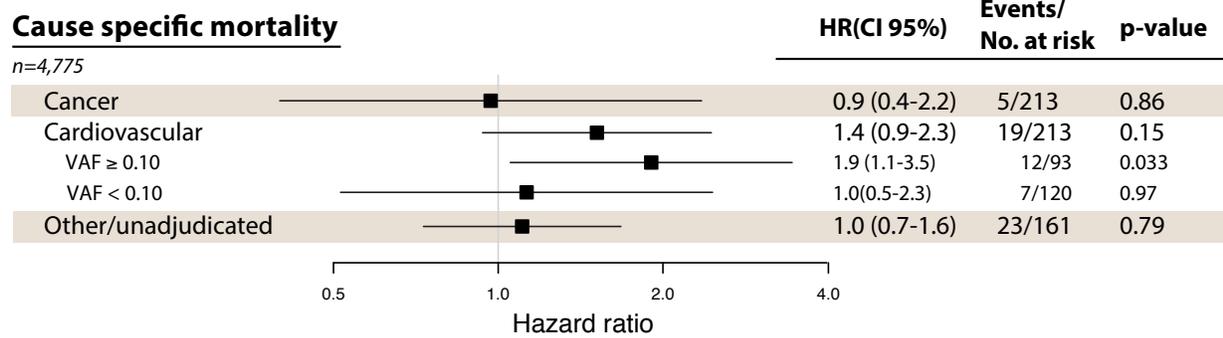


Supplementary Figure S9
Kaplan-Meier Curves for overall survival by cohorts



Supplementary Figure S10 Cause-specific mortality associated with mutations

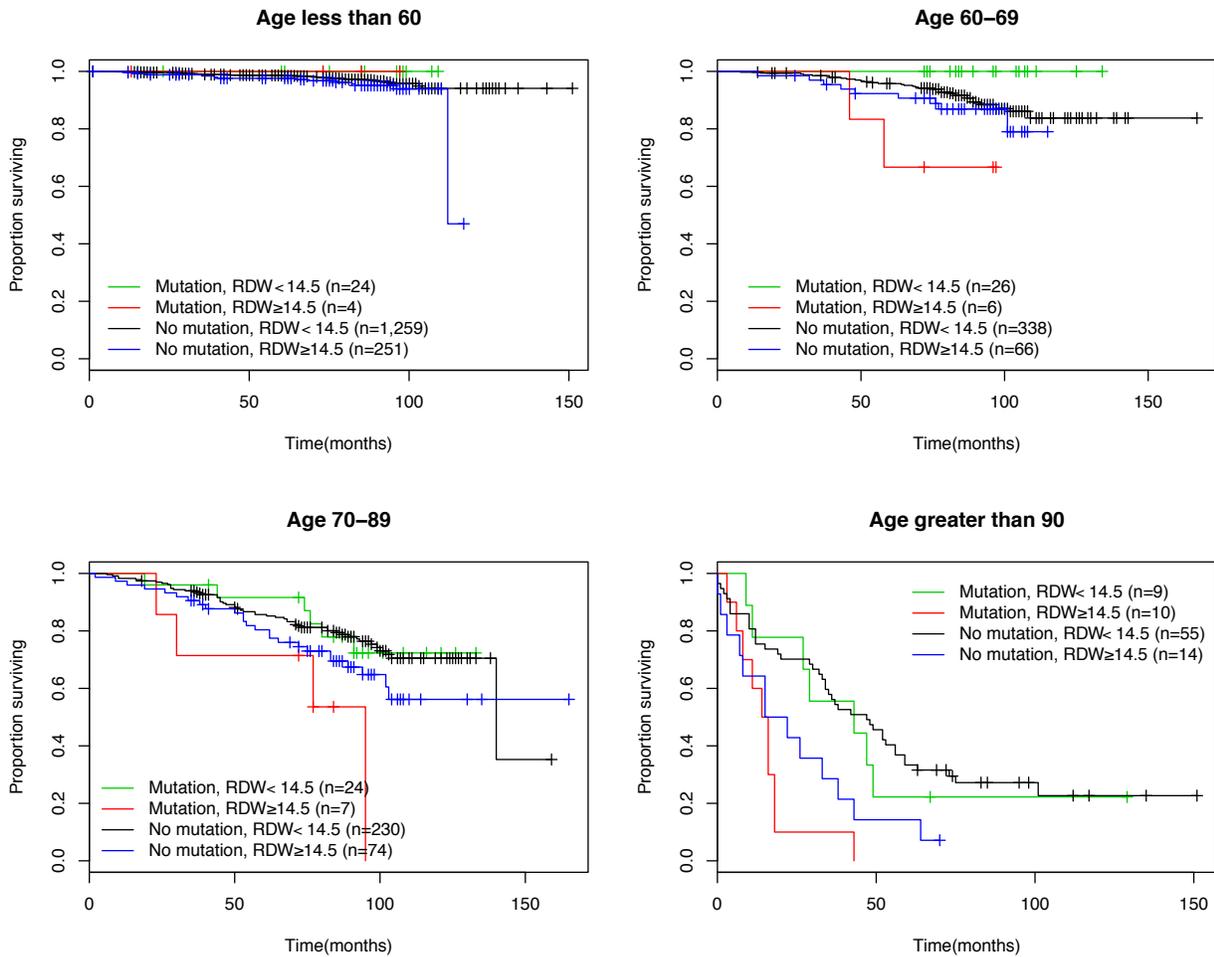
Hazard ratios obtained by competing risks regression, with death by other causes as the competing risk. Results shown are risk associated with clonality. For cancer deaths, only non-hematologic cancers were included. Cardiovascular deaths included fatal strokes (hemorrhagic or ischemic) and fatal myocardial infarction. All regressions included age groups (less than 50, 50-59, 60-69, 70 or older), diabetes status, and sex as covariates. Individuals were from Botnia, FUSION, MEC, and Jackson Heart Study. For Jackson Heart Study, only cardiovascular outcomes were adjudicated (all other deaths were considered unadjudicated).



Supplementary Figure S11

Survival curves for individuals with clonal hematopoiesis

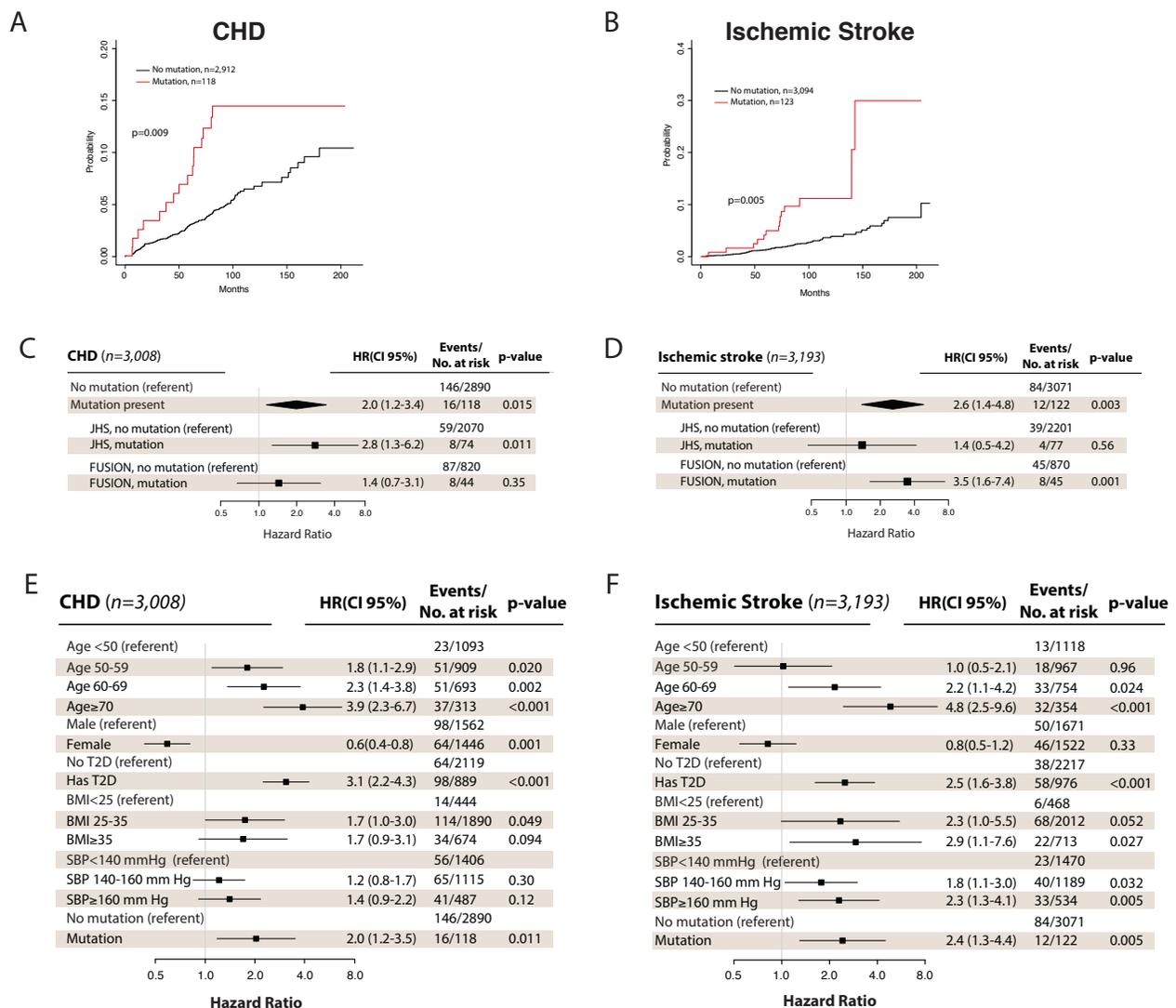
Kaplan-Meier curves for individuals with or without clones, stratified by high ($\geq 14.5\%$) or normal ($< 14.5\%$) red cell distribution width. Hash marks represent censored individuals. Individuals were from Jackson Heart Study or Longevity Genes Project.



Supplementary Figure S12

Association of somatic mutations with incident cardiovascular disease

A-B) Cumulative incidence plots for incident coronary heart disease (CHD) (A) and ischemic stroke (B). Curves were generated from competing risks data with death as the competing risk. Those with prevalent events were excluded from the analyses. C-D) Forest plots for risk of developing incident CHD (C) and ischemic stroke (D) in those with somatic mutations. Diamond represents the results of fixed-effects meta-analysis using beta-coefficients from competing risks regressions for both cohorts, and horizontal lines are 95 percent confidence intervals. Age groups (less than 50, 50-59, 60-69, and 70 or greater), T2D status, sex, systolic blood pressure groups (less than 140 mm Hg, 140-160 mm Hg, and greater than 160 mm Hg), and body mass index groups (less than 25, 25-35, and greater than 35) were included as categorical covariates in the regression models, with death as the competing risk. Those with prevalent events were excluded from the analyses. E-F) Full regression parameters for forest plots from C-D, with Fine-Gray p-values from competing risks regression shown. FUSION and JHS data are combined into a single regression model.



The GoT2D Consortium

Gonçalo Abecasis¹, Vineeta Agarwala², Peter Algren³, David Altshuler^{2,4,5,6,7,8,9}, Martin Hrabé de Angelis¹⁰, Eric Banks², Richard Bergman¹¹, Thomas Blackwell¹, Michael Boehnke¹, Lori Bonnycastle¹², David Buck¹³, Martijn van de Bunt¹⁴, Noël Burt², Peter Chines¹², Francis Collins¹², Mark DePristo², Peter Donnelly¹³, Timothy Fennell², Jason Flannick^{2,5}, Pierre Fontanillas², Timothy Frayling¹⁵, Christian Fuchsberger¹, Stacey Gabriel², Kyle Gaulton¹³, Christian Gieger¹⁰, Harald Grallert¹⁰, Todd Green², Leif Groop³, Christopher Hartl², Andrew Hattersley¹⁵, Bryan Howie¹⁶, Cornelia Huth¹⁰, Jeroen Huyghe¹, Bo Isomaa^{17,18}, Anne Jackson¹⁹, Goo Jun¹, Hyun Min Kang¹, Jasmina Kravic³, Jennifer Kriebel¹⁰, Ashish Kumar¹³, Phoenix Kwan¹, Claes Ladenvall³, Cecilia Lindgren¹³, Adam Locke¹, Gerton Lunter¹³, Clement Ma¹, Anubha Mahajan¹³, Alisa Manning², Mark McCarthy^{13,14,20}, Gil McVean¹³, Christa Meisinger¹⁰, Thomas Meitinger²¹, Karen Mohlke²², Andrew Morris^{13,23}, Loukas Moutsianas¹³, Martina Müller-Nurasyid¹⁰, Pål Njølstad^{24,25}, Richard Pearson¹³, John Perry¹³, Annette Peters¹⁰, Ryan Poplin², Inga Prokopenko^{13,26}, Wolfgang Rathmann¹⁰, William Rayner¹³, Janina Ried¹⁰, Manuel Rivas¹³, Neil Robertson¹³, Laura Scott¹, Khalid Shakir², Xueling Sim¹, Kerrin Small²⁷, Timothy Spector²⁷, Michael Stitzel²⁸, Konstantin Strauch¹⁰, Heather Stringham¹, Tim Strom²¹, Adrian Tan¹, Tanya Teslovich¹, Tiinamajja Toumi^{17,18}, Jaakko Tuomilehto²⁹

The T2D-GENES Consortium

Gonçalo Abecasis¹, Marcio Almeida³⁰, David Altshuler^{2,4,5,6,7,8,9}, Jennifer Asimit³¹, Gil Atzmon³², Mathew Barber¹⁶, Nir Barzilai³², Nicola Beer¹⁴, Graeme Bell^{16,33}, Jennifer Below³⁴, Tom Blackwell¹, John Blangero³⁰, Michael Boehnke¹, Donald Bowden³⁵, Noël Burt², John Chambers^{36,37,38}, Han Chen³⁹, Peng Chen⁴⁰, Peter Hines¹², Sungkyoung Choi⁴¹, Claire Churchhouse², Pablo Cingolani⁴², Belinda Cornes⁴³, Nancy Cox^{16,33}, Aaron Day-Williams³¹, Ravindranath Duggirala³⁰, Josée Dupuis³⁹, Thomas Dyer³⁰, Shuang Feng¹, Juan Fernandez-Tajes¹³, Teresa Ferreira¹³, Tasha Fingerlin⁴⁴, Jason Flannick^{2,5}, Jose Florez^{2,5,6}, Pierre Fontanillas², Timothy Frayling¹⁵, Christian Fuchsberger¹, Eric Gamazon³³, Kyle Gaulton¹³, Saurabh Ghosh⁴⁵, Benjamin Glaser⁴⁶, Anna Gloyn¹⁴, Robert Grossman^{33,47}, Jason Grundstad⁴⁷, Craig Hanis³⁴, Allison Heath⁴⁷, Heather Highland³⁴, Momoko Hirokoshi¹³, Ik-Soo Huh⁴¹, Jeroen Huyghe¹, Kamran Ikram^{48,43,49,50}, Kathleen Jablonski⁵¹, Young Jin Kim⁵², Goo Jun²⁵, Norihiro Kato⁵³, Jayoun Kim⁴¹, Ryan King⁵⁴, Jaspal Kooner^{37,38,55}, Min-Seok Kwon⁴¹, Hae Kyung Im⁵⁴, Markku Laakso⁵⁶, Kevin Koi-Yau Lam⁴⁰, Jaehoon Lee⁴¹, Selyeong Lee⁴¹, Sungyoung Lee⁴⁷, Donna Lehman⁵⁷, Heng Li², Cecilia Lindgren¹³, Xuanyao Liu^{40,58}, Oren Livne¹⁶, Adam Locke¹, Anubha Mahajan¹³, Julian Maller^{13,59}, Alisa Manning², Taylor Maxwell³⁴, Alexander Mazur⁶⁰, Mark McCarthy^{13,14,20}, James Meigs^{6,61}, Byungju Min⁴¹, Karen Mohlke²², Andrew Morris^{13,23}, Solomon Musani⁶², Yoshihiko Nagai⁶⁰, Maggie Ng³⁵, Dan Nicolae^{16,33,12}, Sohee Oh⁴¹, Nicholette Palmer³⁵, Taesung Park⁴¹, Toni Pollin⁶³, Inga Prokopenko^{13,26}, David Reich^{2,4}, Manuel Rivas¹³, Laura Scott¹, Mark Seielstad⁶⁴, Yoon Shin Cho⁶⁵, E-Shyong Tai^{48,40,66}, Xueling Sim¹, Robert Sladek^{60,67}, Philip Smith⁶⁸, Ioanna Tachmazidou³¹, Tanya Teslovich¹, Jason Torres^{16,33}, Vasily Trubetskoy^{16,33}, Sara Willems⁶⁹, Amy Williams^{2,4}, James Wilson⁷⁰, Steven Wiltshire⁷¹, Sungho Won⁷², Andrew Wood¹⁵, Wang Xu⁶⁶, Yik Ying Teo^{73,74,75,58,76}, Joon Yoon⁴¹, Jong-Young Lee⁷⁷, Matthew Zawistowski¹, Eleftheria Zeggini³¹, Weihua Zhang³⁶, Sebastian Zöllner^{1,78}

Affiliations

1. Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA

2. Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA
3. Department of Clinical Sciences, Diabetes and Endocrinology, Clinical Research Centre, Lund University, Malmö, Sweden
4. Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA
5. Center for Human Genetic Research and Diabetes Research Center (Diabetes Unit), Massachusetts General Hospital, Boston, MA, USA
6. Department of Medicine, Harvard Medical School, Boston, MA, USA
7. Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA
8. Department of Molecular Biology, Harvard Medical School, Boston, MA, USA
9. Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA
10. Institute of Genetic Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany
11. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA
12. National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA
13. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
14. Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK
15. Genetics of Complex Traits, University of Exeter Medical School, Exeter, UK
16. Department of Human Genetics, University of Chicago, Chicago, IL, USA
17. Department of Medicine, Helsinki University Hospital, Helsinki, Finland
18. Folkhälsan Research Center, Helsinki, Finland
19. Department of Pathology, University of Michigan, Ann Arbor, MI, USA
20. Oxford NIHR Biomedical Research Centre, Churchill Hospital, Headington, UK
21. Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany
22. Department of Genetics, University of North Carolina-Chapel Hill, Chapel Hill, NC, USA
23. Department of Biostatistics, University of Liverpool, Liverpool, UK
24. Department of Clinical Science, University of Bergen, Bergen, Norway
25. Department of Pediatrics, Haukeland University Hospital, Bergen, Norway
26. Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden
27. Kings College of London, Department of Twin Research and Genetic Epidemiology, London, UK
28. Department of Biochemistry and Molecular Biology, Pennsylvania State University, State College, PA

29. Department of Public Health, University of Helsinki, Helsinki, Finland
30. Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX, USA
31. Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK
32. Department of Medicine, Department of Genetics, Albert Einstein College of Medicine, NY, USA
33. Department of Medicine, University of Chicago, Chicago, IL, USA
34. Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX, USA
35. Center for Genomics and Personalized Medicine Research, Center for Diabetes Research, Department of Biochemistry, Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA
36. Department of Epidemiology and Biostatistics, Imperial College London, London, UK
37. Imperial College Healthcare NHS Trust, London, UK
38. Ealing Hospital National Health Service (NHS) Trust, Middlesex, UK
39. Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA
40. Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore
41. Seoul National University, Seoul, South Korea
42. McGill Centre for Bioinformatics, McGill University, Montr al, Quebec, Canada
43. Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore
44. Department of Epidemiology, Colorado School of Public Health, Aurora, CO, USA
45. Department of Textile Technology, Indian Institute of Technology Delhi, New Delhi, India
46. Department of Endocrinology, Hadassah Medical Center, Kiryat Hadassah, Jerusalem, Israel
47. Institute for Genomics and Systems Biology, University of Chicago, Chicago, IL, USA
48. Duke-National University of Singapore Graduate Medical School, Singapore, Singapore
49. Department of Ophthalmology, National University of Singapore and National University Health System, Singapore, Singapore
50. Department of Ophthalmology, Erasmus Medical Center, Rotterdam, the Netherlands
51. The Biostatistics Center, George Washington University, Rockville, MD, USA
52. Department of Neurology, Konkuk University School of Medicine, Seoul, South Korea
53. Department of Gene Diagnostics and Therapeutics, Research Institute, National Center for Global Health and Medicine, Tokyo, Japan
54. Department of Health Studies, University of Chicago, Chicago, IL, USA
55. National Heart and Lung Institute (NHLI), Imperial College London, Hammersmith Hospital, London, UK

56. Department of Medicine, University of Eastern Finland, Kuopio Campus and Kuopio University Hospital, Kuopio, Finland
57. Division of Clinical Epidemiology, Department of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA
58. Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore, Singapore
59. Department of Statistics, University of Oxford, Oxford, UK
60. McGill University, Montréal, Québec, Canada
61. General Medicine Division, Massachusetts General Hospital, Boston, MA, USA
62. Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA
63. Department of Medicine, Program for Personalized and Genomic Medicine, University of Maryland School of Medicine, Baltimore, MD, USA
64. University of California San Francisco, San Francisco, CA, USA
65. Department of Biomedical Science, Hallym University, Chuncheon, Gangwon-do, South Korea
66. Department of Medicine, National University of Singapore, National University Health System, Singapore, Singapore
67. Department of Medicine, Royal Victoria Hospital, Montréal, Québec, Canada
68. National Institute of Diabetes and Digestive and Kidney Disease, National Institutes of Health, Bethesda, MD, USA
69. Department of Genetic Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands
70. Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS, USA
71. Centre for Medical Research, Western Australian Institute for Medical Research, The University of Western Australia, Nedlands, WA, Australia
72. Chung-Ang University, Seoul, South Korea
73. Department of Epidemiology and Public Health, National University of Singapore, Singapore, Singapore
74. Centre for Molecular Epidemiology, National University of Singapore, Singapore, Singapore
75. Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, Singapore
76. Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore
77. Center for Genome Science, Korea National Institute of Health, Osong Health Technology Administration Complex, Chungcheongbuk-do, South Korea
78. Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA

The SIGMA (Slim Initiative in Genomic Medicine for the Americas) Type 2 Diabetes Genetics Consortium

Writing Team:

Amy L. Williams^{1,2}, Suzanne B. R. Jacobs¹, Hortensia Moreno-Macías³, Alicia Huerta-Chagoya^{4,5}, Claire Churchouse¹, Carla Márquez-Luna⁶, Humberto García-Ortíz⁶, María José Gómez-Vázquez^{4,7}, Noël P. Burt¹, Carlos A. Aguilar-Salinas⁴, Clicerio González-Villalpando^{8,*}, Jose C. Florez^{1,9,10,*}, Lorena Orozco^{6,*}, Christopher A. Haiman^{11,*}, Teresa Tusié-Luna^{4,5,*}, and David Altshuler^{1,2,9,10,12,13,14,*} (* denotes co-corresponding authors)

Analysis Team:

Amy L. Williams^{1,2}, Carla Márquez-Luna⁶, Alicia Huerta-Chagoya^{4,5}, Stephan Ripke^{1,15}, María José Gómez-Vázquez^{4,7}, Alisa K. Manning¹, Hortensia Moreno-Macías³, Humberto García-Ortíz⁶, Benjamin Neale^{1,15}, Noël P. Burt¹, Carlos A. Aguilar-Salinas⁴, David Reich^{1,2}, Daniel O. Stram¹¹, Juan Carlos Fernández-López⁶, Sandra Romero-Hidalgo⁶, David Altshuler^{1,2,9,10,12,13,14}, Jose C. Florez^{1,8,10}, Teresa Tusié-Luna^{4,5}, Nick Patterson¹, and Christopher A. Haiman¹¹

Clinical Research, Study Design and Metabolic Phenotyping:

Diabetes in Mexico Study: Irma Aguilar-Delfín⁶, Angélica Martínez-Hernández⁶, Federico Centeno-Cruz⁶, Elvia Mendoza-Caamal⁶, Cristina Revilla-Monsalve¹⁶, Sergio Islas-Andrade¹⁶, Emilio Córdova⁶, Eunice Rodríguez-Arellano¹⁷, Xavier Soberón⁶, and Lorena Orozco⁶

Massachusetts General Hospital: Jose C. Florez^{1,9,10}

Mexico City Diabetes Study: Clicerio González-Villalpando⁸ and María Elena González-Villalpando⁸

Multiethnic Cohort Study: Christopher A. Haiman¹¹, Brian E. Henderson¹¹, Kristine Monroe¹¹, Lynne Wilkens¹⁸, Laurence N. Kolonel¹⁸, and Loic Le Marchand¹⁸

UNAM/INCMNSZ Diabetes Study: Laura Riba⁵, María Luisa Ordóñez-Sánchez⁴, Rosario Rodríguez-Guillén⁴, Ivette Cruz-Bautista⁴, Maribel Rodríguez-Torres⁴, Linda Liliana Muñoz-Hernández⁴, Tamara Sáenz⁴, Donají Gómez⁴, and Ulices Alvirde⁴

Sample OC and Whole Genome Genotyping:

Noël P. Burt¹, Robert C. Onofrio¹⁹, Wendy M. Brodeur¹⁹, Diane Gage¹⁹, Jacquelyn Murphy¹, Jennifer Franklin¹⁹, Scott Mahan¹⁹, Kristin Ardlie¹⁹, Andrew T. Crenshaw¹⁹, and Wendy Winckler¹⁹

Neandertal analysis team:

Kay Prüfer²⁰, Michael V. Shunkov²¹, Susanna Sawyer²⁰, Udo Stenzel²⁰, Janet Kelso²⁰, Monkol Lek^{1,15}, Sriram Sankararaman^{1,2}, Amy L. Williams^{1,2}, Nick Patterson¹, Daniel G. MacArthur^{1,15}, David Reich^{1,2}, Anatoli P. Derevianko²¹, and Svante Pääbo²⁰

Functional analysis and metabolite profiling:

Suzanne B. R. Jacobs¹, Claire Churchouse¹, Shuba Gopal²², James A. Grammatikos²², Ian C. Smith²³, Kevin H. Bullock²², Amy A. Deik²², Amanda L. Souza²², Kerry A. Pierce²², Clary B. Clish²², and David Altshuler^{1,2,9,10,12,13,14}

Replication genotyping and analysis:

Broad Institute of Harvard and MIT: Timothy Fennell¹⁹, Yossi Farjoun¹⁹, Broad Genomics Platform¹⁹, and Stacey Gabriel¹⁹

Singapore Chinese Health Study: Daniel O. Stram¹¹, Myron D. Gross²⁴, Mark A. Pereira²⁴, Mark Seielstad²⁵, Woon-Puay Koh^{26,27}, and E-Shyong Tai^{26,27,28}

T2D-GENES Consortium: Jason Flannick^{1,9}, Pierre Fontanillas¹, Andrew Morris²⁹, Tanya M. Teslovich³⁰, Noël P. Burt¹, Gil Atzmon³¹, John Blangero³², Don W. Bowden³³, John Chambers^{34,35,36}, Yoon Shin Cho³⁷, Ravindranath Duggirala³², Benjamin Glaser^{38,39}, Craig Hanis⁴⁰, Jaspal Kooner^{35,36,41}, Markku Laakso⁴², Jong-Young Lee⁴³, E-Shyong Tai^{26,27,28}, Yik Ying Teo^{44,45,46,47,48}, James G. Wilson⁴⁹, and the T2D-GENES Consortium

Multiethnic Cohort Study: Christopher A. Haiman¹¹, Brian E. Henderson¹¹, Kristine Monroe¹¹, Lynne Wilkens¹⁸, Laurence N. Kolonel¹⁸, and Loic Le Marchand¹⁸

Texas Biomedical Research Institute and University of Texas Health Science Center at San Antonio: Sobha Puppala³², Vidya S. Farook³², Farook Thameem⁵⁰, Hanna E. Abboud⁵⁰, Ralph A. DeFronzo⁵¹, Christopher P. Jenkinson⁵¹, Donna M. Lehman⁵², Joanne E. Curran³², John Blangero³², and Ravindranath Duggirala³²

Scientific and Project Management:

Noël P. Burt¹ and Maria L. Cortes⁵³

Steering Committee:

David Altshuler^{1,2,9,10,12,13,14}, Jose C. Florez^{1,9,10}, Christopher A. Haiman¹¹, Brian E. Henderson¹¹, Carlos A. Aguilar-Salinas⁴, Clicerio González-Villalpando⁸, Lorena Orozco⁶, and Teresa Tusié-Luna^{4,5}

Table of Affiliations

1 Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA

2 Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA

3 Universidad Autonoma Metropolitana, Mexico City, Mexico

4 Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico

5 Instituto de Investigaciones Biomédicas, UNAM. Unidad de Biología Molecular y Medicina Genómica, UNAM/INCMNSZ, Mexico City, Mexico

6 Instituto Nacional de Medicina Genómica, Mexico City, Mexico

7 Universidad Autónoma de Nuevo León, San Nicolás de los Garza, Nuevo León 66451, México

8 Centro de Estudios en Diabetes, Unidad de Investigación en Diabetes y Riesgo Cardiovascular, Centro de Investigación en Salud Poblacional, Instituto Nacional de Salud Pública, Mexico City, Mexico

9 Center for Human Genetic Research and Diabetes Research Center (Diabetes Unit), Massachusetts General Hospital, Boston, Massachusetts, USA

10 Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA

11 Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California,, USA

12 Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA

13 Department of Molecular Biology, Harvard Medical School, Boston, Massachusetts, USA

14 Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

15 Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA

16 Instituto Mexicano del Seguro Social SXXI, Mexico City, Mexico

17 Instituto de Seguridad y Servicios Sociales para los Trabajadores del Estado, Mexico City, Mexico

18 Epidemiology Program, University of Hawaii Cancer Center, Honolulu, Hawaii, USA

- 19 The Genomics Platform, The Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA
- 20 Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany
- 21 Palaeolithic Department, Institute of Archaeology and Ethnography, Russian Academy of Sciences, Siberian Branch, 630090 Novosibirsk, Russia
- 22 The Metabolite Profiling Platform, The Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA
- 23 Cancer Biology Program, The Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA
- 24 University of Minnesota, Minneapolis, Minnesota, USA
- 25 University of California San Francisco, San Francisco, California, USA
- 26 Duke-National University of Singapore Graduate Medical School, Singapore, Singapore
- 27 Saw Swee Hock School of Public Health, National University of Singapore, Singapore
- 28 Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
- 29 Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
- 30 Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan, USA
- 31 Department of Medicine, Department of Genetics, Albert Einstein College of Medicine, Bronx, New York, USA
- 32 Department of Genetics, Texas Biomedical Research Institute, San Antonio, Texas, USA
- 33 Center for Genomics and Personalized Medicine Research, Center for Diabetes Research, Department of Biochemistry, Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA
- 34 Department of Epidemiology and Biostatistics, Imperial College London, London, UK
- 35 Imperial College Healthcare NHS Trust, London, UK.
- 36 Ealing Hospital National Health Service (NHS) Trust, Middlesex, UK
- 37 Department of Biomedical Science, Hallym University, Chuncheon, Gangwon-do, Korea
- 38 Endocrinology and Metabolism Service, Hadassah-Hebrew University Medical School, Jerusalem, Israel

- 39 Israel Diabetes Research Group (IDRG), Israel
- 40 Human Genetics Center, University of Texas Health Science Center at Houston, Houston, Texas, USA
- 41 National Heart and Lung Institute (NHLI), Imperial College London, Hammersmith Hospital, London, UK
- 42 Department of Medicine, University of Eastern Finland, Kuopio Campus and Kuopio University Hospital, Kuopio, Finland
- 43 Center for Genome Science, National Institute of Health, Osong Health Technology Administration Complex, Chungcheongbuk-do, Cheongwon-gun, Gangoe-myeon, Yeonje-ri, Korea
- 44 Department of Epidemiology and Public Health, National University of Singapore, Singapore, Singapore
- 45 Centre for Molecular Epidemiology, National University of Singapore, Singapore, Singapore
- 46 Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, Singapore
- 47 Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore, Singapore
- 48 Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore.
- 49 Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, Mississippi, USA
- 50 Division of Nephrology, Department of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, Texas
- 51 Division of Diabetes, Department of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, Texas
- 52 Division of Clinical Epidemiology, Department of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, Texas
- 53 Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA